```
OM protein - nucleic search, using frame_plus_p2n model

Run on:          October  6, 2004, 18:18:04 ; Search time 833 Seconds
                                              (without alignments)
                                              4584.790 Million cell updates/sec

Title:           US-09-970-944-2
Perfect score:   4787
Sequence:        1 MAVRPGLWPALLGIVLAAWL..........AVAGLGQPDAGLFTVSEAEC 899

Scoring table:   BLOSUM62
                 Xgapop 10.0 , Xgapext  0.5
                 Ygapop 10.0 , Ygapext  0.5
                 Fgapop  6.0 , Fgapext  7.0
                 Delop   6.0 , Delext   7.0

Searched:        3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters:      6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09970944/runat_05102004_112009_1616/app_query.fasta_1.1095
-DB=N_Geneseq_29Jan04 -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09970944_@CGN_1_1_581_@runat_05102004_112009_1616 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :       N_Geneseq_29Jan04:*
                 1:  geneseqn1980s:*
                 2:  geneseqn1990s:*
                 3:  geneseqn2000s:*
                 4:  geneseqn2001as:*
                 5:  geneseqn2001bs:*
                 6:  geneseqn2002s:*
                 7:  geneseqn2003as:*
                 8:  geneseqn2003bs:*
                 9:  geneseqn2003cs:*
                 10:  geneseqn2004s:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 4787 | 100.0 | 2881 | 6 | ABK49422 | Abk49422 DNA encod |
| 2 | 4698.5 | 98.2 | 2752 | 6 | ABK37922 | Abk37922 cDNA enco |
| 3 | 4688.5 | 97.9 | 2697 | 6 | ABK52891 | Abk52891 Human net |
| 4 | 4545.5 | 95.0 | 2697 | 6 | AAS16843 | Aas16843 Rat netri |
| 5 | 4545.5 | 95.0 | 3014 | 2 | AAV52940 | Aav52940 Rat UNC-5 |
| 6 | 4434 | 92.6 | 2907 | 4 | AAK52261 | Aak52261 Human pol |
| 7 | 4379.5 | 91.5 | 3580 | 6 | ABK15169 | Abk15169 Human REP |
| 8 | 2694 | 56.3 | 1787 | 2 | AAV52941 | Aav52941 Human UNC |
| 9 | 2673.5 | 55.8 | 3646 | 5 | AAS75738 | Aas75738 DNA encod |
| 10 | 2490.5 | 52.0 | 2995 | 6 | ABK92062 | Abk92062 DNA encod |
| 11 | 2484 | 51.9 | 2860 | 6 | ABT06279 | Abt06279 Human NOV |
| 12 | 2484 | 51.9 | 2860 | 6 | ABT06280 | Abt06280 Human NOV |
| 13 | 2479 | 51.8 | 3884 | 4 | AAS21316 | Aas21316 Human cDN |
| 14 | 2479 | 51.8 | 3884 | 7 | ACD23925 | Acd23925 Novel hum |
| 15 | 2479 | 51.8 | 3884 | 7 | ACA67066 | Aca67066 cDNA enco |
| 16 | 2479 | 51.8 | 3884 | 7 | ACA03675 | Aca03675 cDNA enco |
| 17 | 2479 | 51.8 | 3884 | 7 | ABX89213 | Abx89213 DNA encod |
| 18 | 2479 | 51.8 | 3884 | 7 | ACD41867 | Acd41867 Human sec |
| 19 | 2479 | 51.8 | 3884 | 7 | ACA04096 | Aca04096 Human cDN |
| 20 | 2479 | 51.8 | 3884 | 8 | ADA45664 | Ada45664 Novel hum |
| 21 | 2479 | 51.8 | 3884 | 8 | ADA76095 | Ada76095 Human PRO |
| 22 | 2479 | 51.8 | 3884 | 8 | ADA18745 | Ada18745 Human PRO |
| 23 | 2479 | 51.8 | 3884 | 8 | ADA61368 | Ada61368 Homo sapi |
| 24 | 2479 | 51.8 | 3884 | 8 | ADB19153 | Adb19153 Novel hum |
| 25 | 2479 | 51.8 | 3884 | 8 | ADB27694 | Adb27694 cDNA enco |
| 26 | 2479 | 51.8 | 3884 | 8 | ADA86173 | Ada86173 Novel hum |
| 27 | 2479 | 51.8 | 3884 | 8 | ADB15737 | Adb15737 Human PRO |
| 28 | 2479 | 51.8 | 3884 | 8 | ADA47523 | Ada47523 Human PRO |
| 29 | 2479 | 51.8 | 3884 | 8 | ADA67318 | Ada67318 Human PRO |
| 30 | 2479 | 51.8 | 3884 | 8 | ADB30325 | Adb30325 cDNA enco |
| 31 | 2479 | 51.8 | 3884 | 8 | ADA85621 | Ada85621 Novel hum |
| 32 | 2479 | 51.8 | 3884 | 8 | ADA96833 | Ada96833 Human PRO |
| 33 | 2479 | 51.8 | 3884 | 8 | ADA79137 | Ada79137 Human PRO |
| 34 | 2479 | 51.8 | 3884 | 8 | ADA87276 | Ada87276 Novel hum |
| 35 | 2479 | 51.8 | 3884 | 8 | ADB16478 | Adb16478 Human PRO |
| 36 | 2479 | 51.8 | 3884 | 8 | ADA91570 | Ada91570 Novel hum |
| 37 | 2479 | 51.8 | 3884 | 8 | ADB14633 | Adb14633 Human PRO |
| 38 | 2479 | 51.8 | 3884 | 8 | ADB18594 | Adb18594 Novel hum |
| 39 | 2479 | 51.8 | 3884 | 8 | ADA93809 | Ada93809 Human PRO |
| 40 | 2479 | 51.8 | 3884 | 8 | ADB19705 | Adb19705 Novel hum |
| 41 | 2479 | 51.8 | 3884 | 8 | ADB13017 | Adb13017 Human PRO |
| 42 | 2479 | 51.8 | 3884 | 8 | ACD98496 | Acd98496 Novel hum |
| 43 | 2479 | 51.8 | 3884 | 8 | ADA74271 | Ada74271 Human PRO |
| 44 | 2479 | 51.8 | 3884 | 8 | ADB24504 | Adb24504 Human PRO |
| 45 | 2479 | 51.8 | 3884 | 8 | ADA82028 | Ada82028 Human PRO |

ALIGNMENTS

```
RESULT 1
ABK49422
ID   ABK49422 standard; DNA; 2881 BP.
XX
AC   ABK49422;
XX
DT   15-JUL-2002  (first entry)
XX
DE   DNA encoding human UNC5-like protein NOV1.
XX
KW   Human; NOVX polypeptide; cardiomyopathy; atherosclerosis; cancer;
KW   cell signal processing; metabolic pathway modulation; cancerous tissue;
KW   antibody; diabetes; transgenic animal; UNC5-like protein; NOV1;
KW   chromosome 13; gene; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             87. .2786
FT                   /*tag=  a
FT                   /product= "Human UNC5-like protein NOV1"
XX
PN   WO200229038-A2.
XX
PD   11-APR-2002.
XX
PF   04-OCT-2001; 2001WO-US031377.
XX
PR   04-OCT-2000; 2000US-0237862P.
XX
PA   (CURA-) CURAGEN CORP.
XX
```

```
PI   Herrmann JL,  Rastelli L,  Shimkets RA;
XX
DR   WPI; 2002-340104/37.
DR   P-PSDB; AAU79939.
XX
PT   Novel isolated NOVX polypeptide, and encoded polynucleotide, useful for
PT   treating cardiomyopathy, artherosclerosis, and cancer.
XX
PS   Claim 8; Page 7-8; 180pp; English.
XX
CC   The present invention relates to a new NOVX polypeptide having a 900
CC   (NOV1), 4349 (NOV2), 940 (NOV3), 798 (NOV4), 865 (NOV5), or 331 (NOV6)
CC   residue amino acid sequence, as given in the specification. The novel
CC   polypeptide, and its encoding polynucleotide, are used to treat
CC   cardiomyopathy, atherosclerosis, cancer or a disease related to cell
CC   signal processing and metabolic pathway modulation, in a human. Detecting
CC   the polypeptide or polynucleotide is useful for identifying cancerous
CC   tissue. The antibody can be used to treat diabetes or cancer. The host
CC   cells can be used to produce non-human transgenic animals useful in drug
CC   screening. The present nucleic acid sequence is that of the human UNC5-
CC   like NOV1 gene located on chromosome 13. This sequence encodes the human
CC   UNC5-like protein NOV1 of the invention
XX
SQ   Sequence 2881 BP; 526 A; 985 C; 868 G; 502 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              0            Length:          2881
Score:                  4787.00      Matches:         899
Percent Similarity:     100.00%      Conservative:    0
Best Local Similarity:  100.00%      Mismatches:      0
Query Match:            100.00%      Indels:          0
DB:                     6            Gaps:            0

US-09-970-944-2 (1-899) x ABK49422 (1-2881)

Qy          1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         87 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 146

Qy         21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 206

Qy         41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 266

Qy         61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 326

Qy         81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCG 386

Qy        101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 446

Qy        121 GluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 506

Qy        141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 566

Qy        161 SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 626

Qy        181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 686

Qy        201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 746

Qy        221 CysValAlaLysAsnIleValAlaArgArgArgSerAlaSerAlaAlaValIleValTyr 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        747 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 806

Qy        241 ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        807 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 866

Qy        261 GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        867 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 926

Qy        281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        927 TGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGC 986

Qy        301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        987 AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC 1046

Qy        321 ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1047 CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG 1106

Qy        341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1107 GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG 1166

Qy        361 AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuLeuValLeu 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1167 GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC 1226

Qy        381 IleLeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIle 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1227 ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1286

Qy        401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1287 CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1346

Qy        421 LeuThrIleGlnProAspLeuSerThrThrThrThrTyrGlnGlySerLeuCysProArg 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1347 CTCACCATCCAGCCGGACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGG 1406

Qy        441 GlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeuGly 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1407 CAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGT 1466

Qy        461 GlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheValSer 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1467 GGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCC 1526

Qy        481 ArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThrTyr 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1527 CGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT 1586

Qy        501 GlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeuLeu 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1587 GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTC 1646

Qy        521 IleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHisLys 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1647 ATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG 1706

Qy        541 ProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleValSer 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1707 CCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC 1766
```

```
Qy        561 CysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGly 580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1767 TGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG 1826

Qy        581 GluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySerTrp 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1827 GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG 1886

Qy        601 GluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGlnLeu 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1887 GAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTG 1946

Qy        621 GluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGlu 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1947 GAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAG 2006

Qy        641 AlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAlaCys 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2007 GCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGC 2066

Qy        661 ThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuLys 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2067 ACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAG 2126

Qy        681 GluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgValLeu 700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2127 GAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTG 2186

Qy        701 HisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSerLeu 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2187 CACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTG 2246

Qy        721 TrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsnGly 740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2247 TGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGC 2306

Qy        741 ThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSerAsp 760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2307 ACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGAC 2366

Qy        761 LeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIleAsn 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2367 CTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAAC 2426

Qy        781 PheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAlaGly 800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2427 TTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGG 2486

Qy        801 ValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLysIle 820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2487 GTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATA 2546

Qy        821 IleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGlnLys 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2547 ATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAA 2606

Qy        841 LeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMetIle 860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2607 CTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATC 2666

Qy        861 LeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAlaAla 880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2667 CTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCA 2726

Qy        881 ValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2727 GTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2783

RESULT 2
ABK37922
ID   ABK37922 standard; cDNA; 2752 BP.
XX
AC   ABK37922;
XX
DT   21-MAY-2002  (first entry)
XX
DE   cDNA encoding Human protein NOV1.
XX
KW   Human; NOVX; ss; gene; cardiomyopathy; atherosclerosis; diabetes;
KW   cell signal processing disorder; metabolic disorder; obesity; infection;
KW   anorexia; cancer-associated cachexia; cancer; neurodegenerative disorder;
KW   Alzheimer's disease; Parkinson's disease; immune disorder;
KW   haematopoietic disorders; dyslipidaemia; pain; asthma; hypertension;
KW   osteoporosis; Crohn's disease; multiple sclerosis; angina pectoris;
KW   myocardial infarction; ulcer; allergy; benign prostatic hypertrophy;
KW   psychosis; neurological disorder; anxiety; schizophrenia;
KW   manic depression; dementia; dyskinesia; Huntington's disease;
KW   Gilles de la Tourette's syndrome; gene therapy.
XX
OS   Homo sapiens.
XX
PN   WO200210216-A2.
XX
PD   07-FEB-2002.
XX
PF   30-JUL-2001; 2001WO-US024225.
XX
PR   28-JUL-2000; 2000US-0221409P.
PR   04-AUG-2000; 2000US-0222840P.
PR   04-AUG-2000; 2000US-0223752P.
PR   04-AUG-2000; 2000US-0223762P.
PR   04-AUG-2000; 2000US-0223769P.
PR   04-AUG-2000; 2000US-0223770P.
PR   14-AUG-2000; 2000US-0225146P.
PR   15-AUG-2000; 2000US-0225392P.
PR   15-AUG-2000; 2000US-0225470P.
PR   16-AUG-2000; 2000US-0225697P.
PR   01-FEB-2001; 2001US-0263662P.
PR   05-APR-2001; 2001US-0281645P.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Padigaru M,  Mezes P,  Mishra V,  Burgess C,  Casman S,  Grosse WM;
PI   Alsobrook JP,  Lepley DM,  Gerlach VL,  Macdougall JR,  Smithson G;
XX
DR   WPI; 2002-180074/23.
DR   P-PSDB; AAU85403.
XX
PT   New isolated cytoplasmic, nuclear, membrane bound, or secreted
PT   polypeptide, useful for treating cardiomyopathy, atherosclerosis,
PT   infections, cancer, neurodegenerative, metabolic, hematopoietic and
PT   immune disorders.
XX
PS   Claim 9; Page 9-10; 213pp; English.
XX
CC   The invention relates to an isolated cytoplasmic, nuclear, membrane
CC   bound, or secreted polypeptide (NOVX, x= 1-14) their variants or mature
CC   form. Also included are the nucleic acids encoding the NOVX proteins, a
CC   vector comprising the nucleic acid, a cell comprising the vector, an anti
CC   -NOVX antibody and modulators of NOVX. NOVX, the nucleic acid and the
CC   antibody are useful for treating or preventing a NOVX-associated
CC   disorder, where the disorder is selected from cardiomyopathy,
CC   atherosclerosis, diabetes, a disorder related to cell signal processing
CC   and metabolic pathway modulation, metabolic disorders, obesity,
CC   infectious disease, anorexia, cancer-associated cachexia, cancer,
CC   neurodegenerative disorders, Alzheimer's disease, Parkinson's disease,
CC   immune disorders, haematopoietic disorders, and the various
CC   dyslipidaemias, metabolic disturbances associated with obesity, the
CC   metabolic syndrome X and wasting disorders associated with chronic
CC   diseases, bacterial, fungal, protozoal and viral infections, pain,
CC   bulimia, asthma, hypertension, urinary retention, osteoporosis, Crohn's
CC   disease, multiple sclerosis, Albright Hereditary Osteodystrophy, angina
CC   pectoris, myocardial infarction, ulcer, allergy, benign prostatic
CC   hypertrophy, and psychotic and neurological disorders, including anxiety,
CC   schizophrenia, manic depression, delirium, dementia, and dyskinesias,
```

```
CC   such as Huntington's disease and Gilles de la Tourette's syndrome. The
CC   nucleic acid is useful in gene therapy. The present sequence encodes a
CC   NOVX protein
XX
SQ   Sequence 2752 BP; 505 A; 937 C; 829 G; 481 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              9.21e-315      Length:          2752
Score:                  4698.50        Matches:         888
Percent Similarity:     98.89%         Conservative:    2
Best Local Similarity:  98.67%         Mismatches:      7
Query Match:            98.15%         Indels:          3
DB:                     6              Gaps:            3

US-09-970-944-2 (1-899) x ABK37922 (1-2752)

Qy      1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeu 20
Db     46 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 105

Qy     21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40
Db    106 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 165

Qy     41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro 60
Db    166 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 225

Qy     61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80
Db    226 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 285

Qy     81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100
Db    286 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 345

Qy    101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120
Db    346 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 405

Qy    121 GluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAla 140
Db    406 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 465

Qy    141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160
Db    466 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 525

Qy    161 SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu 180
Db    526 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 585

Qy    181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200
Db    586 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 645

Qy    201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220
Db    646 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 705

Qy    221 CysValAlaLysAsnIleValAlaArgArgArgSerAlaSerAlaAlaValIleValTyr 240
Db    706 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 765

Qy    241 ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg 260
Db    766 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 825

Qy    261 GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe 280
Db    826 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 885

Qy    281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300
Db    886 TGTGAGGGGCAGAATGTC---CAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGC 942

Qy    301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320
Db    943 AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC 1002

Qy    321 ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeu 340
Db   1003 CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG 1062

Qy    341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360
Db   1063 GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG 1122

Qy    361 AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuValLeu 380
Db   1123 GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC 1182

Qy    381 IleLeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIle 400
Db   1183 ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1242

Qy    401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420
Db   1243 CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1302

Qy    421 LeuThrIleGlnProAspLeuSer---ThrThrThrThrTyrGlnGlySerLeuCysPro 439
Db   1303 CTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCC 1362

Qy    440 ArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu 459
Db   1363 CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG 1422

Qy    460 GlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheVal 479
Db   1423 GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC 1482

Qy    480 SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr 499
Db   1483 TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC 1542

Qy    500 TyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeu 519
Db   1543 TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC 1602

Qy    520 LeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHis 539
Db   1603 CTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC 1662

Qy    540 LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal 559
Db   1663 AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT 1722

Qy    560 SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys 579
Db   1723 AGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT 1782

Qy    580 GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySer 599
Db   1783 GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC 1842

Qy    600 TrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGln 619
Db   1843 TGGGAG---GATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG 1899

Qy    620 LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly 639
Db   1900 CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA 1959

Qy    640 GluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAla 659
```

```
Db       1960 GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC 2019

Qy        660 CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu 679
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2020 TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC 2079

Qy        680 LysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgVal 699
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2080 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTC 2139

Qy        700 LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2140 CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCC 2199

Qy        720 LeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn 739
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2200 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2259

Qy        740 GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2260 GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2319

Qy        760 AspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIle 779
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2320 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2379

Qy        780 AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAla 799
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2380 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2439

Qy        800 GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2440 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2499

Qy        820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2500 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2559

Qy        840 LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2560 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG 2619

Qy        860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2620 ATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2679

Qy        880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2680 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2739
```

RESULT 3
ABK52891

```
ID   ABK52891 standard; DNA; 2697 BP.
XX
AC   ABK52891;
XX
DT   27-AUG-2002  (first entry)
XX
DE   Human netrin binding membrane receptor UNC5H-1 DNA sequence #1.
XX
KW   Netrin binding membrane receptor; receptor; UNC5H-1; gene; ds; human;
KW   nootropic; neuroprotective; cytostatic; antiparkinsonian;
KW   cerebroprotective; cancer; central nervous system; CNS; stroke;
KW   Parkinson's disease; multiple sclerosis; Alzheimer's disease.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .2697
FT                   /*tag=  a
FT                   /product= "Netrin binding membrane receptor UNC5H-1"
XX
PN   WO200233080-A2.
XX
PD   25-APR-2002.
XX
PF   15-OCT-2001; 2001WO-EP011891.
XX
PR   16-OCT-2000; 2000US-0240061P.
XX
PA   (FARB ) BAYER AG.
XX
PI   Koehler RH;
XX
DR   WPI; 2002-463314/49.
DR   P-PSDB; AAU97899.
XX
PT   Novel human netrin binding membrane receptor polypeptide and
PT   polynucleotides for identifying modulating agents useful in treating
PT   diseases e.g. Parkinson's disease, multiple sclerosis, stroke,
PT   Alzheimer's disease.
XX
PS   Claim 1; Fig 1; 94pp; English.
XX
CC   This invention relates to the DNA and protein sequences of a novel
CC   purified human netrin binding membrane receptor, UNC5H-1. The DNA
CC   sequence of the invention is useful as a probe for detecting a nucleic
CC   acid encoding the UNC5H-1 protein in a biological sample. The sequences
CC   of the invention are useful to screen for agents which decrease the
CC   activity of the UNC5H-1 protein. The sequences are also useful for
CC   screening agents which regulate (modulate) the activity of the protein of
CC   the invention. A pharmaceutical composition containing the protein of the
CC   invention or a reagent that modulates the activity of the UNC5H-1 protein
CC   may be useful for treating a UNC5H-1 dysfunction related disease such as
CC   cancer or a central nervous system (CNS) disorders (e.g, Parkinson's
CC   disease, multiple sclerosis, stroke and Alzheimer's disease). Fusion
CC   proteins comprising the UNC5H-1 protein are useful for generating
CC   antibodies and for in various assay systems, and the protein can be used
CC   as a bait protein in a two-hybrid assay or three-hybrid assay. The method
CC   of the invention is useful for detecting a coding sequence for the UNC5H-
CC   1 protein. The present sequence represents a DNA sequence encoding the
CC   human netrin binding membrane receptor UNC5H-1 protein of the invention
XX
SQ   Sequence 2697 BP; 503 A; 906 C; 807 G; 481 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:               4.4e-314       Length:          2697
Score:                   4688.50        Matches:         886
Percent Similarity:      98.78%         Conservative:    3
Best Local Similarity:   98.44%         Mismatches:      8
Query Match:             97.94%         Indels:          3
DB:                      6              Gaps:            3

US-09-970-944-2 (1-899) x ABK52891 (1-2697)

Qy          1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 60

Qy         21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120

Qy         41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 180

Qy         61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 240

Qy         81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100
              |||||||||||||||||||||||||||||||||||||||||||||||||||    |||
Db        241 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGCTGCCC 300

Qy        101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120
```

```
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 360

Qy        121 GluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 420

Qy        141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160
              ||||||||||||||    ||||||||||||||||||||||||||||||||||||||||||
Db        421 TACATCCGCATAGCCTATTTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 480

Qy        161 SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 540

Qy        181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 600

Qy        201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 660

Qy        221 CysValAlaLysAsnIleValAlaArgArgArgSerAlaSerAlaAlaValIleValTyr 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 720

Qy        241 ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 780

Qy        261 GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 840

Qy        281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300
              ||||||||||||||||||         :::|||       :::   |||     ||||||||
Db        841 TGTGAGGGGCAGAATGTC---CAGAAAACAGCCTGCGCCACCCTGTGCCCAGTGGACGGC 897

Qy        301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        898 AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC 957

Qy        321 ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        958 CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG 1017

Qy        341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360
              |||||||||||||||||||||||||||||||||:::||||||||||||||||||||||||
Db       1018 GACACCCGCAACTGTACCAGTGACCTCTGTGTACACACTGCTTCTGGCCCTGAGGACGTG 1077

Qy        361 AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuLeuValLeu 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1078 GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC 1137

Qy        381 IleLeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIle 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1138 ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1197

Qy        401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1198 CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1257

Qy        421 LeuThrIleGlnProAspLeuSer---ThrThrThrThrTyrGlnGlySerLeuCysPro 439
              ||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
Db       1258 CTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCC 1317

Qy        440 ArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu 459
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1318 CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG 1377

Qy        460 GlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheVal 479
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1378 GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC 1437

Qy        480 SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr 499
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1438 TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC 1497

Qy        500 TyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeu 519
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1498 TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGAATCAGCCTC 1557

Qy        520 LeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHis 539
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1558 CTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC 1617

Qy        540 LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal 559
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1618 AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT 1677

Qy        560 SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys 579
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1678 AGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT 1737

Qy        580 GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySer 599
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1738 GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC 1797

Qy        600 TrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGln 619
              ||||||      ||||||||||||||||||||||||||||||||||||||||||||||||
Db       1798 TGGGAG---GATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG 1854

Qy        620 LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly 639
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1855 CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA 1914

Qy        640 GluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAla 659
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1915 GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC 1974

Qy        660 CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu 679
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1975 TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC 2034

Qy        680 LysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgVal 699
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2035 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTC 2094

Qy        700 LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2095 CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCC 2154

Qy        720 LeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn 739
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2155 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2214

Qy        740 GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2215 GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2274

Qy        760 AspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIle 779
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2275 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2334

Qy        780 AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAla 799
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2335 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2394

Qy        800 GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2395 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2454

Qy        820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2455 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2514
```

```
Qy        840 LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859
Db       2515 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG 2574

Qy        860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879
Db       2575 ATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2634

Qy        880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
Db       2635 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2694

RESULT 4
AAS16843
ID   AAS16843 standard; cDNA; 2697 BP.
XX
AC   AAS16843;
XX
DT   14-FEB-2002  (first entry)
XX
DE   Rat netrin receptor UNC5H1 (YSG7) cDNA.
XX
KW   YSG; YSG7; schizophrenia; chronic animal model; LCGU; netrin receptor;
KW   local cerebral glucose utilisation; phosphodiesterase 1-alpha; UNC5H1;
KW   calcium-independent alpha-latrotoxin receptor; CIRL; trkE; synapsin 1A;
KW   epithelial discoidin domain receptor 1; synapsin 1B; neuroleptic; ss;
KW   tumour necrosis factor alpha; TNF-alpha; rat.
XX
OS   Rattus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .2697
FT                   /*tag=  a
FT                   /product= "Rat netrin receptor UNC5H1"
XX
PN   WO200175440-A2.
XX
PD   11-OCT-2001.
XX
PF   02-APR-2001; 2001WO-GB001486.
XX
PR   31-MAR-2000; 2000GB-00007880.
PR   26-MAY-2000; 2000GB-00012768.
XX
PA   (WELF-) WELFIDE CORP.
XX
PI   Cochran S,  Paterson G,  Ohashi Y,  Morris B,  Pratt J;
XX
DR   WPI; 2002-010813/01.
DR   P-PSDB; AAU10543.
XX
PT   Novel chronic animal model of schizophrenia, useful for identifying anti-
PT   psychotic drugs and genes that are associated with schizophrenia.
XX
PS   Claim 1; Fig 8a; 79pp; English.
XX
CC   The invention relates to YSG polynucleotide fragments for use in
CC   diagnosing and/or developing treatments for schizophrenia using chronic
CC   animal models. The polynucleotides and their encoded polypeptides are
CC   used for identification of compounds which modulate the expression of YSG
CC   molecules, leading to the manufacture of schizophrenia medicaments. The
CC   sequences can also be used for testing candidate compounds for any effect
CC   on the polypeptides. Anti-schizophrenic effects of a compound can be
CC   determined by measuring local cerebral glucose utilisation (LCGU) or
CC   comparing its expression level with that of a control group. The
CC   sequences are useful in the identification of genes associated with
CC   schizophrenic states and in the development of an antibody. The sequences
CC   of the invention include phosphodiesterase 1-alpha, calcium-independent
CC   alpha-latrotoxin receptors (CIRL)-1,2&3, epithelial discoidin domain
CC   receptor 1 (trkE), netrin receptor (UNC5H1), synapsins 1A and AB and
CC   tumour necrosis factor (TNF) alpha. This sequence represents rat netrin
CC   receptor UNC5H1 (YSG7) DNA
XX
SQ   Sequence 2697 BP; 541 A; 864 C; 766 G; 526 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:               3.32e-304      Length:         2697
Score:                   4545.50        Matches:        852
Percent Similarity:      96.78%         Conservative:   19
Best Local Similarity:   94.67%         Mismatches:     26
Query Match:             94.96%         Indels:         3
DB:                      6              Gaps:           3

US-09-970-944-2 (1-899) x AAS16843 (1-2697)

Qy          1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeu 20
Db          1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGTGCTCCTGGGCATAGTCCTCGCCGCCTGGCTT 60

Qy         21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40
Db         61 CGTGGTTCGGGTGCCCAGCAGAGTGCCACGGTGGCCAATCCAGTGCCCGGTGCCAACCCC 120

Qy         41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro 60
Db        121 GACCTGCTGCCCCACTTCCTGGTAGAGCCTGAGGACGTGTACATTGTCAAGAACAAGCCG 180

Qy         61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80
Db        181 GTGTTGTTGGTGTGCAAGGCTGTGCCTGCCACCCAGATCTTCTTCAAGTGCAATGGGGAA 240

Qy         81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100
Db        241 TGGGTCCGCCAGGTCGATCACGTAATTGAACGCAGCACCGACAGCAGCAGCGGATTGCCA 300

Qy        101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120
Db        301 ACCATGGAGGTCCGTATCAACGTATCGAGGCAGCAGGTAGAGAAAGTGTTTGGGCTGGAG 360

Qy        121 GluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAla 140
Db        361 GAATACTGGTGCCAGTGTGTGGCATGGAGCTCCTCGGGTACCACCAAAAGTCAGAAGGCC 420

Qy        141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160
Db        421 TACATCCGGATTGCCTATTTGCGCAAGAACTTTGAGCAGGAGCCACTGGCCAAGGAAGTG 480

Qy        161 SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu 180
Db        481 TCACTGGAGCAAGGCATTGTACTACCTTGTCGCCCCCCAGAAGGAATCCCCCCAGCTGAG 540

Qy        181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200
Db        541 GTGGAGTGGCTTCGAAATGAGGACCTCGTGGACCCCTCCCTCGATCCCAATGTGTACATC 600

Qy        201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220
Db        601 ACGCGGGAGCACAGCCTAGTCGTGCGTCAGGCCCGCCTGGCCGACACGGCCAACTACACC 660

Qy        221 CysValAlaLysAsnIleValAlaArgArgArgSerAlaSerAlaAlaValIleValTyr 240
Db        661 TGTGTGGCCAAGAACATCGTAGCCCGTCGCCGAAGCACCTCTGCAGCGGTCATTGTTTAT 720

Qy        241 ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg 260
Db        721 GTGAACGGTGGGTGGTCGACGTGGACTGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGT 780

Qy        261 GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe 280
Db        781 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCACCTCTCAACGGGGGCGCCTTC 840

Qy        281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300
Db        841 TGTGAGGGGCAGAATGTC---CAGAAAACAGCCTGCGCCACTCTGTGCCCAGTGGATGGG 897
```

```
Qy      301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320
Db      898 AGCTGGAGTTCGTGGAGTAAGTGGTCAGCCTGTGGGCTTGACTGCACCCACTGGCGGAGC 957

Qy      321 ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeu 340
Db      958 CGCGAGTGCTCTGACCCAGCACCCCGCAATGGAGGTGAGGAGTGTCGGGGTGCTGACCTG 1017

Qy      341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360
Db     1018 GACACCCGCAACTGTACCAGTGACCTCTGCCTGCACACCGCTTCTTGCCCCGAGGACGTG 1077

Qy      361 AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuLeuValLeu 380
Db     1078 GCTCTCTACATCGGCCTTGTCGCTGTGGCTGTGTGCCTCTTCTTGCTGTTGCTGGCCCTT 1137

Qy      381 IleLeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIle 400
Db     1138 GGACTCATTTACTGTCGCAAGAAGGAAGGGCTGGACTCCGATGTGGCCGACTCGTCCATC 1197

Qy      401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420
Db     1198 CTCACCTCGGGCTTCCAGCCTGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCACCTG 1257

Qy      421 LeuThrIleGlnProAspLeuSer---ThrThrThrThrTyrGlnGlySerLeuCysPro 439
Db     1258 CTCACCATCCAGCCAGACCTCAGCACCACCACTACCACCTACCAGGGCAGTCTATGTTCG 1317

Qy      440 ArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu 459
Db     1318 AGGCAGGATGGACCCAGCCCCAAGTTCCAGCTCTCTAATGGTCACCTGCTCAGCCCACTG 1377

Qy      460 GlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheVal 479
Db     1378 GGGAGTGGCCGCCATACGTTGCACCACAGCTCACCCACCTCTGAGGCTGAGGACTTCGTC 1437

Qy      480 SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr 499
Db     1438 TCCCGCCTCTCCACCCAAAACTACTTTCGTTCCCTGCCCCGCGGCACCAGCAACATGGCC 1497

Qy      500 TyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeu 519
Db     1498 TACGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACGGGGATCAGCCTC 1557

Qy      520 LeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHis 539
Db     1558 CTCATACCCCCGGATGCCATCCCCCGAGGAAAGATCTACGAGATCTACCTCACACTGCAC 1617

Qy      540 LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal 559
Db     1618 AAGCCAGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCAGTCGTT 1677

Qy      560 SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys 579
Db     1678 AGCTGTGGGCCCCCAGGAGTCCTGCTCACCCGGCCAGTCATCCTTGCAATGGACCACTGT 1737

Qy      580 GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySer 599
Db     1738 GGAGAGCCCAGCCCTGACAGCTGGAGTCTGCGCCTCAAAAAGCAGTCCTGCGAGGGCAGT 1797

Qy      600 TrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGln 619
Db     1798 TGGGAG---GATGTGCTGCACCTTGGTGAGGAGTCACCTTCCCACCTCTACTACTGCCAG 1854

Qy      620 LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly 639
Db     1855 CTGGAGGCCGGGGCCTGCTATGTCTTCACGGAGCAGCTGGGCCGCTTTGCCCTGGTAGGA 1914

Qy      640 GluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAla 659
Db     1915 GAGGCCCTCAGCGTGGCTGCCACCAAGCGCCTCAGGCTCCTTCTGTTTGCTCCCGTGGCC 1974

Qy      660 CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu 679
Db     1975 TGTACGTCCCTTGAGTACAACATCCGAGTGTACTGCCTACACGACACCCACGACGCTCTC 2034

Qy      680 LysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgVal 699
Db     2035 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTAGGTGGACAGCTGATCCAGGAGCCTCGCGTC 2094

Qy      700 LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719
Db     2095 CTGCACTTCAAAGACAGTTACCACAACCTACGTCTCTCCATCCACGACGTGCCCAGCTCC 2154

Qy      720 LeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn 739
Db     2155 CTGTGGAAGAGCAAGCTACTTGTCAGCTACCAGGAGATCCCTTTTTACCACATCTGGAAC 2214

Qy      740 GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759
Db     2215 GGCACCCAGCAGTATCTGCACTGCACCTTCACCCTGGAGCGCATCAACGCCAGCACCAGC 2274

Qy      760 AspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIle 779
Db     2275 GACCTGGCCTGCAAGGTGTGGGTGTGGCAGGTGGAGGGAGATGGGCAGAGCTTCAACATC 2334

Qy      780 AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAla 799
Db     2335 AACTTCAACATCACTAAGGACACAAGGTTTGCTGAATTGTTGGCTCTGGAGAGTGAAGGG 2394

Qy      800 GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
Db     2395 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAAAAG 2454

Qy      820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839
Db     2455 ATCATCGCCAGTCTGGACCCACCCTGCAGCCGGGGCGCCGACTGGAGAACTCTAGCCCAG 2514

Qy      840 LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859
Db     2515 AAACTTCACCTGGACAGCCATCTTAGCTTCTTTGCCTCCAAGCCCAGCCCTACAGCCATG 2574

Qy      860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879
Db     2575 ATCCTCAACCTATGGAGGCACGGCACTTCCCCAACGGCAACCTCGGCCAGCTGGCAGCA 2634

Qy      880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
Db     2635 GCTGTGGCCGGACTGGGCCAACCAGATGCTGGCCTCTTCACGGTGTCGGAGGCCGAGTGT 2694
```

```
RESULT 5
AAV52940
ID   AAV52940 standard; cDNA; 3014 BP.
XX
AC   AAV52940;
XX
DT   25-MAR-2003  (revised)
DT   21-DEC-1998  (first entry)
XX
DE   Rat UNC-5 homologue unc5h-1 cDNA.
XX
KW   UNC-5; UNC5H-1; rat; netrin receptor; cell migration; axon guidance;
KW   diagnosis; therapy; ds.
XX
OS   Rattus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..2697
FT                   /*tag=  a
XX
PN   WO9837085-A1.
XX
PD   27-AUG-1998.
XX
PF   19-FEB-1998;   98WO-US003143.
XX
```

```
PR   19-FEB-1997;   97US-00808982.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Tessier-Lavigne M,  Leonardo ED,  Hinck L,  Masu M,  Keinomasu K;
XX
DR   WPI; 1998-495364/42.
DR   P-PSDB; AAW78898.
XX
PT   Netrin-binding, vertebrate proteins - useful for diagnosis, therapy and
PT   the biopharmaceutical industry.
XX
PS   Claim 7; Page 15-17; 32pp; English.
XX
CC   This cDNA, termed unc5h-1, comprises a rat homologue of Caenorhabditis
CC   elegans unc-5. Rat unc5h-1 and unc5h-2 (see AAV52942) cDNAs were isolated
CC   from an E18 brain cDNA library. The predicted proteins (see AAW78898 and
CC   AAW78900) show similarity with UNC-5. They are predicted to be involved
CC   in cell migration and axon guidance, and are characterised as receptor
CC   proteins for netrins. Gene expression is observed in regions where
CC   differentiating neurons are undergoing axogogenesis. Human unc5h-1 (see
CC   AAV52941) and unc5h-2 (see AAV52943) cDNAs are also claimed. Vertebrate
CC   UNC-5 proteins may be produced recombinantly from transfected host cells
CC   by utilising these vertebrate UNC-5 nucleic acids. The invention also
CC   provides unc-5 hybridisation probes and primers, vertebrate UNC-5-
CC   specific binding agents such as specific antibodies, and methods of
CC   making and using the subject compositions in diagnosis (e.g. genetic
CC   hybridisation screens for vertebrate unc-5 transcripts), therapy (e.g.
CC   gene therapy to modulate vertebrate unc-5 gene expression) and in the
CC   biopharmaceutical industry (e.g. as immunogens, reagents for modulating
CC   cell guidance, reagents for screening chemical libraries for lead
CC   pharmacological agents, etc.). (Updated on 25-MAR-2003 to correct PI
CC   field.)
XX
SQ   Sequence 3014 BP; 596 A; 977 C; 849 G; 592 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              3.84e-304      Length:         3014
Score:                  4545.50        Matches:        852
Percent Similarity:     96.78%         Conservative:   19
Best Local Similarity:  94.67%         Mismatches:    26
Query Match:            94.96%         Indels:         3
DB:                     2              Gaps:           3

US-09-970-944-2 (1-899) x AAV52940 (1-3014)

Qy          1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeu 20
Db          1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGTGCTCCTGGGCATAGTCCTCGCCGCCTGGCTT 60

Qy         21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40
Db         61 CGTGGTTCGGGTGCCCAGCAGAGTGCCACGGTGGCCAATCCAGTGCCCGGTGCCAACCCC 120

Qy         41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro 60
Db        121 GACCTGCTGCCCCACTTCCTGGTAGAGCCTGAGGACGTGTACATTGTCAAGAACAAGCCG 180

Qy         61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80
Db        181 GTGTTGTTGGTGTGCAAGGCTGTGCCTGCCACCCAGATCTTCTTCAAGTGCAATGGGGAA 240

Qy         81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100
Db        241 TGGGTCCGCCAGGTCGATCACGTAATTGAACGCAGCACCGACAGCAGCAGCGGATTGCCA 300

Qy        101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120
Db        301 ACCATGGAGGTCCGTATCAACGTATCGAGGCAGCAGGTAGAGAAAGTGTTTGGGCTGGAG 360

Qy        121 GluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAla 140
Db        361 GAATACTGGTGCCAGTGTGTGGCATGGAGCTCCTCGGGTACCACCAAAAGTCAGAAGGCC 420

Qy        141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160
Db        421 TACATCCGGATTGCCTATTTGCGCAAGAACTTTGAGCAGGAGCCACTGGCCAAGGAAGTG 480

Qy        161 SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu 180
Db        481 TCACTGGAGCAAGGCATTGTACTACCTTGTCGCCCCCCAGAAGGAATCCCCCCAGCTGAG 540

Qy        181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200
Db        541 GTGGAGTGGCTTCGAAATGAGGACCTCGTGGACCCCTCCCTCGATCCCAATGTGTACATC 600

Qy        201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220
Db        601 ACGCGGGAGCACAGCCTAGTCGTGCGTCAGGCCCGCCTGGCCGACACGGCCAACTACACC 660

Qy        221 CysValAlaLysAsnIleValAlaArgArgArgSerAlaSerAlaAlaValIleValTyr 240
Db        661 TGTGTGGCCAAGAACATCGTAGCCCGTCGCCGAAGCACCTCTGCAGCGGTCATTGTTTAT 720

Qy        241 ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg 260
Db        721 GTGAACGGTGGGTGGTCGACGTGGACTGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGT 780

Qy        261 GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe 280
Db        781 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCACCTCTCAACGGGGGCGCCTTC 840

Qy        281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300
Db        841 TGTGAGGGGCAGAATGTC---CAGAAAACAGCCTGCGCCACTCTGTGCCCAGTGGATGGG 897

Qy        301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320
Db        898 AGCTGGAGTTCGTGGAGTAAGTGGTCAGCCTGTGGGCTTGACTGCACCCACTGGCGGAGC 957

Qy        321 ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeu 340
Db        958 CGCGAGTGCTCTGACCCAGCACCCCGCAATGGAGGTGAGGAGTGTCGGGGTGCTGACCTG 1017

Qy        341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360
Db       1018 GACACCCGCAACTGTACCAGTGACCTCTGCCTGCACACCGCTTCTTGCCCCGAGGACGTG 1077

Qy        361 AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuLeuValLeu 380
Db       1078 GCTCTCTACATCGGCCTTGTCGCTGTGGCTGTGTGCCTCTTCTTGCTGTTGCTGGCCCTT 1137

Qy        381 IleLeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIle 400
Db       1138 GGACTCATTTACTGTCGCAAGAAGGAAGGGCTGGACTCCGATGTGGCCGACTCGTCCATC 1197

Qy        401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420
Db       1198 CTCACCTCGGGCTTCCAGCCTGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCACCTG 1257

Qy        421 LeuThrIleGlnProAspLeuSer---ThrThrThrThrTyrGlnGlySerLeuCysPro 439
Db       1258 CTCACCATCCAGCCAGACCTCAGCACCACCACTACCACCTACCAGGGCAGTCTATGTTCG 1317

Qy        440 ArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu 459
Db       1318 AGGCAGGATGGACCCAGCCCCAAGTTCCAGCTCTCTAATGGTCACCTGCTCAGCCCACTG 1377

Qy        460 GlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheVal 479
Db       1378 GGGAGTGGCCGCCATACGTTGCACCACAGCTCACCCACCTCTGAGGCTGAGGACTTCGTC 1437

Qy        480 SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr 499
Db       1438 TCCCGCCTCTCCACCCAAAACTACTTTCGTTCCCTGCCCCGCGGCACCAGCAACATGGCC 1497
```

```
Qy        500 TyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeu 519
Db       1498 TACGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACGGGGATCAGCCTC 1557

Qy        520 LeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHis 539
Db       1558 CTCATACCCCCGGATGCCATCCCCCGAGGAAAGATCTACGAGATCTACCTCACACTGCAC 1617

Qy        540 LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal 559
Db       1618 AAGCCAGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCAGTCGTT 1677

Qy        560 SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys 579
Db       1678 AGCTGTGGGCCCCCAGGAGTCCTGCTCACCCGGCCAGTCATCCTTGCAATGGACCACTGT 1737

Qy        580 GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySer 599
Db       1738 GGAGAGCCCAGCCCTGACAGCTGGAGTCTGCGCCTCAAAAAGCAGTCCTGCGAGGGCAGT 1797

Qy        600 TrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGln 619
Db       1798 TGGGAG---GATGTGCTGCACCTTGGTGAGGAGTCACCTTCCCACCTCTACTACTGCCAG 1854

Qy        620 LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly 639
Db       1855 CTGGAGGCCGGGGCCTGCTATGTCTTCACGGAGCAGCTGGGCCGCTTTGCCCTGGTAGGA 1914

Qy        640 GluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAla 659
Db       1915 GAGGCCCTCAGCGTGGCTGCCACCAAGCGCCTCAGGCTCCTTCTGTTTGCTCCCGTGGCC 1974

Qy        660 CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu 679
Db       1975 TGTACGTCCCTTGAGTACAACATCCGAGTGTACTGCCTACACGACACCCACGACGCTCTC 2034

Qy        680 LysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgVal 699
Db       2035 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTAGGTGGACAGCTGATCCAGGAGCCTCGCGTC 2094

Qy        700 LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719
Db       2095 CTGCACTTCAAAGACAGTTACCACAACCTACGTCTCTCCATCCACGACGTGCCCAGCTCC 2154

Qy        720 LeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn 739
Db       2155 CTGTGGAAGAGCAAGCTACTTGTCAGCTACCAGGAGATCCCTTTTTACCACATCTGGAAC 2214

Qy        740 GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759
Db       2215 GGCACCCAGCAGTATCTGCACTGCACCTTCACCCTGGAGCGCATCAACGCCAGCACCAGC 2274

Qy        760 AspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIle 779
Db       2275 GACCTGGCCTGCAAGGTGTGGGTGTGGCAGGTGGAGGGAGATGGCAGAGCTTCAACATC 2334

Qy        780 AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAla 799
Db       2335 AACTTCAACATCACTAAGGACACAAGGTTTGCTGAATTGTTGGCTCTGGAGAGTGAAGGG 2394

Qy        800 GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
Db       2395 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAAAAG 2454

Qy        820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839
Db       2455 ATCATCGCCAGTCTGGACCCACCCTGCAGCCGGGGCGCCGACTGGAGAACTCTAGCCCAG 2514

Qy        840 LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859
Db       2515 AAACTTCACCTGGACAGCCATCTTAGCTTCTTTGCCTCCAAGCCCAGCCCTACAGCCATG 2574

Qy        860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879
Db       2575 ATCCTCAACCTATGGGAGGCACGGCACTTCCCCAACGGCAACCTCGGCCAGCTGGCAGCA 2634

Qy        880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
Db       2635 GCTGTGGCCGGACTGGGCCAACCAGATGCTGGCCTCTTCACGGTGTCGGAGGCCGAGTGT 2694

RESULT 6
AAK52261
ID   AAK52261 standard; cDNA; 2907 BP.
XX
AC   AAK52261;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human polynucleotide SEQ ID NO 806.
XX
KW   Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW   vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW   tissue growth factor; immunomodulatory; cancer; leukaemia;
KW   nervous system disorder; arthritis; inflammation; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157190-A2.
XX
PD   09-AUG-2001.
XX
PF   05-FEB-2001; 2001WO-US004098.
XX
PR   03-FEB-2000; 2000US-00496914.
PR   27-APR-2000; 2000US-00560875.
PR   20-JUN-2000; 2000US-00598075.
PR   19-JUL-2000; 2000US-00620325.
PR   01-SEP-2000; 2000US-00654936.
PR   15-SEP-2000; 2000US-00663561.
PR   20-OCT-2000; 2000US-00693325.
PR   30-NOV-2000; 2000US-00728422.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Drmanac RT,  Asundi V,  Zhou P,  Xu C,  Cao Y;
PI   Ma Y,  Zhao QA,  Wang D,  Wang J,  Zhang J,  Ren F,  Chen R,  Wang ZW;
PI   Xue AJ,  Yang Y,  Wejhrman T,  Goodrich R;
XX
DR   WPI; 2001-476283/51.
DR   P-PSDB; AAM79128.
XX
PT   Nucleic acids encoding polypeptides with cytokine-like activities, useful
PT   in diagnosis and gene therapy.
XX
PS   Claim 1; Page 2691-2694; 6221pp; English.
XX
CC   The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC   encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
CC   cytokine, cell proliferation or cell differentiation or which may induce
CC   production of other cytokines in other cell populations. The
CC   polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC   peptide therapy. The polypeptides have various cytokine-like activities,
CC   e.g. stem cell growth factor activity, haematopoiesis regulating
CC   activity, tissue growth factor activity, immunomodulatory activity and
CC   activin/inhibin activity and may be useful in the diagnosis and/or
CC   treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC   inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
CC   (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
CC   sequence listing were missing at the time of publication
XX
SQ   Sequence 2907 BP; 552 A; 966 C; 881 G; 508 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:            1.84e-296      Length:         2907
Score:                4434.00        Matches:        853
Percent Similarity:   90.78%         Conservative:   4
```

```
Best Local Similarity:  90.36%      Mismatches:    17
Query Match:            92.63%      Indels:        70
DB:                     4           Gaps:          7

US-09-970-944-2 (1-899) x AAK52261 (1-2907)

Qy          1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeu 20
              |||      ||||||    |||                          :::|||
Db        148 ATGACGAGGCGTCCCTCCCTG-----------------------ATGGCGGGCAGACAG 183

Qy         21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40
                 |||      ||||||||||||||||||||||||||||||||||||||||||||||
Db        184 CACGGATGGAGCGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 243

Qy         41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 303

Qy         61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 363

Qy         81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100
              |||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        364 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 423

Qy        101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 483

Qy        121 GluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 543

Qy        141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160
              ||||||||||||||   |||||||||||||||||||||||||||||||||||||||||||
Db        544 TACATCCGCATAGCCTATTTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 603

Qy        161 SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        604 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 663

Qy        181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        664 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 723

Qy        201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        724 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 783

Qy        221 CysValAlaLysAsnIleValAlaArgArgArgSerAlaSerAlaAlaValIleValTyr 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        784 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 843

Qy        240 ------------------------------------------------------------ 240

Db        844 GGTGGGCCCCGGGACTCCCTGGTCACAGGGAGAGGCACTGCGGTGCCCCTGGGCAGTGAC 903

Qy        241 ---------------------------ValAsnGlyGlyTrpSerThrTrpThrGluTrp 251
                                         |||||||||||||||||||||||||||||||||
Db        904 ATGTGGCTGTCCTTCTCTGTCCGGCCAGTGAACGGTGGGTGGTCGACGTGGACCGAGTGG 963

Qy        252 SerValCysSerAlaSerCysGlyArgGlyTrpGlnLysArgSerArgSerCysThrAsn 271
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        964 TCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAAC 1023

Qy        272 ProAlaProLeuAsnGlyGlyAlaPheCysGluGlyGlnAsnValHisAspArgThrVal 291
              ||||||||||||||||||||||||||||||||||||||||||||      :::|||
Db       1024 CCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAATGTC---CAGAAAACAGCC 1080

Qy        292 SerSerLeuLeuValSerValAspGlySerTrpSerProTrpSerLysTrpSerAlaCys 311
                 :::   |||         ||||||||||||||||||||||||||||||||||||||
Db       1081 TGCGCCACCCTGTGCCCAGTGGACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGT 1140

Qy        312 GlyLeuAspCysThrHisTrpArgSerArgGluCysSerAspProAlaProArgAsnGly 331
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGA 1200

Qy        332 GlyGluGluCysGlnGlyThrAspLeuAspThrArgAsnCysThrSerAspLeuCysVal 351
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGTACCAGTGACCTCTGTGTA 1260

Qy        352 His---------------------------------------------SerAlaSerGly 356
              |||                                              :::|||||||||
Db       1261 CACAACTCCTACACCCCTGCCCCCACCAAGGCCATGCTGTCTCCCGCAGCTGCTTCTGGC 1320

Qy        357 ProGluAspValAlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeu 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CCTGAGGACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTG 1380

Qy        377 LeuLeuValLeuIleLeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAla 396
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CTGCTTGTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCT 1440

Qy        397 AspSerSerIleLeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAsp 416
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GACTCGTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGAC 1500

Qy        417 AsnProHisLeuLeuThrIleGlnProAspLeuSer---ThrThrThrThrTyrGlnGly 435
              ||||||||||||||||||||||||||||||||||   ||||||||||||||||||||||
Db       1501 AACCCCCATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGC 1560

Qy        436 SerLeuCysProArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeu 455
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 AGTCTCTGTCCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTG 1620

Qy        456 LeuSerProLeuGlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAla 475
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 CTCAGCCCCCTGGGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCC 1680

Qy        476 GluGluPheValSerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThr 495
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 GAGGAGTTCGTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACC 1740

Qy        496 SerAsnMetThrTyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThr 515
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 AGCAACATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACA 1800

Qy        516 GlyIleSerLeuLeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyr 535
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 GGAATCAGCCTCCTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTAC 1860

Qy        536 LeuThrLeuHisLysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeu 555
              ||||||||||||||||||||||||
Db       1861 CTCACGCTGCACAAGCCGGAAGAC------------------------------------ 1884

Qy        556 SerProIleValSerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAla 575
                       |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1885 ---------GTGAGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCT 1935

Qy        576 MetAspHisCysGlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSer 595
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1936 ATGGACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCG 1995

Qy        596 CysGluGlySerTrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeu 615
              ||||||||||||||||||   |||||||||||||||||||||||||||||||||||||||
Db       1996 TGCGAGGGCAGCTGGGAG---GATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTC 2052

Qy        616 TyrTyrCysGlnLeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPhe 635
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2053 TACTACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTT 2112

Qy        636 AlaLeuValGlyGluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPhe 655
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2113 GCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTT 2172
```

```
Qy        656 AlaProValAlaCysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThr 675
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2173 GCGCCGGTGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACC 2232

Qy        676 HisAspAlaLeuLysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGln 695
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2233 CACGATGCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAG 2292

Qy        696 GluProArgValLeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAsp 715
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2293 GAGCCACGGGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGAT 2352

Qy        716 ValProSerSerLeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyr 735
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2353 GTGCCCAGCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTAT 2412

Qy        736 HisIleTrpAsnGlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSer 755
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2413 CACATCTGGAATGGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGC 2472

Qy        756 ProSerThrSerAspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGln 775
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2473 CCCAGCACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAG 2532

Qy        776 SerPheSerIleAsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeu 795
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2533 AGCTTCAGCATCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTG 2592

Qy        796 GluSerGluAlaGlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeu 815
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2593 GAGAGTGAAGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTC 2652

Qy        816 IleArgGlnLysIleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArg 835
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2653 ATTCGGCAGAAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGG 2712

Qy        836 ThrLeuAlaGlnLysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSer 855
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2713 ACTCTGGCCCAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGC 2772

Qy        856 ProThrAlaMetIleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSer 875
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2773 CCCACAGCCATGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGC 2832

Qy        876 GlnLeuAlaAlaAlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSer 895
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2833 CAGCTGGCTGCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCG 2892

Qy        896 GluAlaGluCys 899
              ||||||||||||
Db       2893 GAGGCTGAGTGC 2904
```

RESULT 7
ABK15169
ID   ABK15169 standard; DNA; 3580 BP.
XX
AC   ABK15169;
XX
DT   23-APR-2002  (first entry)
XX
DE   Human REPTR 1 cDNA sequence.
XX
KW   REPTR; human; antiinflammatory; cytostatic; immunosuppressive; antiviral;
KW   anti-HIV; antiarthritic; anticonvulsant; nootropic; neuroprotective;
KW   antiallergic; antibody; immunogen; endometriosis;
KW   gastrointestinal disorder; gastritis; oesophageal carcinoma;
KW   Crohn's disease; irritable bowel syndrome; ulcerative colitis;
KW   endocrine disorder; hypothalamus disorder; Kallman's disease;
KW   autoimmune disease; inflammatory disease; infertility; receptor;
KW   acquired immune deficiency syndrome; AIDS; rheumatoid arthritis; allergy;
KW   osteoarthritis; diabetes mellitus; multiple sclerosis;
KW   systemic lupus erythematosus; cell proliferative disorder; cancer;
KW   developmental disorder; Duchenne muscular dystrophy; gene;
KW   Becker muscular dystrophy; neurological disorder; epilepsy;
KW   Alzheimer's disease; Huntington's disease; reproductive disorder; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             4. .2532
FT                   /*tag=  a
FT                   /product= "REPTR1 protein"
XX
PN   WO200198354-A2.
XX
PD   27-DEC-2001.
XX
PF   21-JUN-2001; 2001WO-US019942.
XX
PR   21-JUN-2000; 2000US-0214027P.
PR   25-AUG-2000; 2000US-0228045P.
PR   12-DEC-2000; 2000US-0255104P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Griffin JA,  Kallick DA,  Tribouley CM,  Yue H,  Nguyen DB,  Tang YT;
PI   Lal P,  Policky JL,  Azimzai Y,  Lu DAM,  Graul R,  Yao MG,  Burford N;
PI   Hafalia AJA,  Baughn MR,  Bandman O,  Patterson C,  Yang J,  Xu Y;
PI   Gandhi AR,  Warren BA,  Ding L,  Sanjanwala MS,  Duggan BM,  Lu Y;
XX
DR   WPI; 2002-090432/12.
DR   P-PSDB; AAU17818.
XX
PT   Twelve human receptors (referred to as REPTR-1 to REPTR-12), useful in
PT   the diagnosis, treatment and prevention of gastrointestinal (e.g.
PT   gastritis), autoimmune/inflammatory (e.g. osteoarthritis) and cell
PT   proliferative (e.g. cancer)disorders.
XX
PS   Claim 57; Page 142-143; 157pp; English.
XX
CC   This invention relates to twelve human receptors cDNA sequences referred
CC   to as REPTR-1 to REPTR-12), and the proteins encoded thereby. The
CC   proteins of the invention may have antiinflammatory, cytostatic,
CC   immunosuppressive, antiviral, anti-HIV, antiarthritic, muscular active
CC   general, anticonvulsant, nootropic, neuroprotective, antiallergic
CC   activities. The sequences of the invention may be used to produce REPTR
CC   agonists or antagonists, and the protein sequences may be used to raise
CC   anti-REPTR antibodies. These molecules and the REPTR polynucleotides and
CC   polypeptides of the invention are useful in the diagnosis,treatment and
CC   prevention of gastrointestinal (e.g. gastritis, oesophageal carcinoma,
CC   Crohn's disease, irritable bowel syndrome, ulcerative colitis), endocrine
CC   (e.g. hypothalamus disorder, Kallman's disease), autoimmune/ inflammatory
CC   (e.g. acquired immune deficiency syndrome (AIDS), rheumatoid arthritis,
CC   allergies, osteoarthritis, diabetes mellitus, multiple sclerosis,
CC   systemic lupus erythematosus), cell proliferative (e.g. cancer),
CC   developmental (e.g. Duchenne and Becker muscular dystrophy), neurological
CC   (e.g. epilepsy, Alzheimer's disease, Huntington's disease) and
CC   reproductive (e.g. infertility, endometriosis) disorders. Numerous other
CC   examples of each disorder are given in the specification. The present
CC   sequence represents the human REPTR1 cDNA sequence of the invention
XX
SQ   Sequence 3580 BP; 670 A; 1233 C; 1025 G; 652 T; 0 U; 0 Other;

```
Alignment Scores:
Pred. No.:              1.41e-292    Length:         3580
Score:                  4379.50      Matches:        838
Percent Similarity:     93.22%       Conservative:   1
Best Local Similarity:  93.11%       Mismatches:     2
Query Match:            91.49%       Indels:         59
DB:                     6            Gaps:           3

US-09-970-944-2 (1-899) x ABK15169 (1-3580)

Qy          1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 63
```

```
Qy    21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40
Db    64 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 123

Qy    41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro 60
Db   124 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 183

Qy    61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80
Db   184 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 243

Qy    81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100
Db   244 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 303

Qy   101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120
Db   304 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 363

Qy   121 GluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAla 140
Db   364 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 423

Qy   141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160
Db   424 TACATCCGCATAGCCTATTTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 483

Qy   161 SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu 180
Db   484 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 543

Qy   181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200
Db   544 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 603

Qy   201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220
Db   604 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 663

Qy   221 CysValAlaLysAsnIleValAlaArgArgArgSerAlaSerAlaAlaValIleValTyr 240
Db   664 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 723

Qy   241 ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg 260
Db   723 ------------------------------------------------------------ 723

Qy   261 GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe 280
Db   723 ------------------------------------------------------------ 723

Qy   281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300
Db   724 ---------------------------------------------------GTGGACGGC 732

Qy   301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320
Db   733 AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC 792

Qy   321 ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeu 340
Db   793 CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG 852

Qy   341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360
Db   853 GACACCCGCAACTGTACCAGTGACCTCTGTGTACACACTGCTTCTGGCCCTGAGGACGTG 912

Qy   361 AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuLeuValLeu 380
Db   913 GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC 972

Qy   381 IleLeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIle 400
Db   973 ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1032

Qy   401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420
Db  1033 CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1092

Qy   421 LeuThrIleGlnProAspLeuSer---ThrThrThrThrTyrGlnGlySerLeuCysPro 439
Db  1093 CTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCC 1152

Qy   440 ArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu 459
Db  1153 CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG 1212

Qy   460 GlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheVal 479
Db  1213 GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC 1272

Qy   480 SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr 499
Db  1273 TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC 1332

Qy   500 TyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeu 519
Db  1333 TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGAATCAGCCTC 1392

Qy   520 LeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHis 539
Db  1393 CTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC 1452

Qy   540 LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal 559
Db  1453 AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT 1512

Qy   560 SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys 579
Db  1513 AGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT 1572

Qy   580 GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySer 599
Db  1573 GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC 1632

Qy   600 TrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGln 619
Db  1633 TGGGAG---GATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG 1689

Qy   620 LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly 639
Db  1690 CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA 1749

Qy   640 GluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAla 659
Db  1750 GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC 1809

Qy   660 CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu 679
Db  1810 TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC 1869

Qy   680 LysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgVal 699
Db  1870 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTC 1929

Qy   700 LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719
Db  1930 CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCC 1989

Qy   720 LeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn 739
Db  1990 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2049

Qy   740 GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2050 GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2109

Qy        760 AspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIle 779
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2110 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2169

Qy        780 AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAla 799
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2170 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2229

Qy        800 GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2230 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2289

Qy        820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2290 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2349

Qy        840 LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2350 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG 2409

Qy        860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2410 ATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2469

Qy        880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2470 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2529

RESULT 8
AAV52941
ID   AAV52941 standard; cDNA; 1787 BP.
XX
AC   AAV52941;
XX
DT   25-MAR-2003  (revised)
DT   21-DEC-1998  (first entry)
XX
DE   Human UNC-5 homologue unc5h-1 cDNA.
XX
KW   UNC-5; UNC5H-1; human; netrin receptor; cell migration; axon guidance;
KW   diagnosis; therapy; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             3. .1670
FT                   /*tag=  a
FT                   /transl_except= (pos:21. .22, aa:Xaa)
FT                   /transl_except= (pos:200. .203, aa:Ile)
FT                   /transl_except= (pos:771. .772, aa:Ser)
FT                   /transl_except= (pos:785. .786, aa:Leu)
FT                   /transl_except= (pos:1078. .1079, aa:Xaa)
FT                   /transl_except= (pos:1098. .1099, aa:Xaa)
FT                   /transl_except= (pos:1106. .1107, aa:Xaa)
FT                   /transl_except= (pos:1621. .1622, aa:Gly)
FT                   /note= "these codons have apparent 1 nucleotide deletions
FT                   or insertions, which alter the reading frame"
XX
PN   WO9837085-A1.
XX
PD   27-AUG-1998.
XX
PF   19-FEB-1998;    98WO-US003143.
XX
PR   19-FEB-1997;    97US-00808982.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Tessier-Lavigne M,  Leonardo ED,  Hinck L,  Masu M,  Keinomasu K;
XX
DR   WPI; 1998-495364/42.
DR   P-PSDB; AAW78899.
XX
PT   Netrin-binding, vertebrate proteins - useful for diagnosis, therapy and
PT   the biopharmaceutical industry.
XX
PS   Claim 7; Page 17; 32pp; English.
XX
CC   This partial cDNA, termed unc5h-1, comprises a human homologue of
CC   Caenorhabditis elegans unc-5. unc5h-1 and unc5h-2 (see AAV52943) cDNAs
CC   were isolated from an embryonic brain cDNA library. The predicted
CC   proteins (see AAW78899 and AAW78901) show similarity with UNC-5. They are
CC   predicted to be involved in cell migration and axon guidance, and are
CC   characterised as receptor proteins for netrins. Gene expression is
CC   observed in regions where differentiating neurons are undergoing
CC   axogogenesis. Rat unc5h-1 (see AAV52940) and unc5h-2 (see AAV52942) cDNAs
CC   are also claimed. Vertebrate UNC-5 proteins may be produced recombinantly
CC   from transfected host cells by utilising these vertebrate UNC-5 nucleic
CC   acids. The invention also provides unc-5 hybridisation probes and
CC   primers, vertebrate UNC-5-specific binding agents such as specific
CC   antibodies, and methods of making and using the subject compositions in
CC   diagnosis (e.g. genetic hybridisation screens for vertebrate unc-5
CC   transcripts), therapy (e.g. gene therapy to modulate vertebrate unc-5
CC   gene expression) and in the biopharmaceutical industry (e.g. as
CC   immunogens, reagents for modulating cell guidance, reagents for screening
CC   chemical libraries for lead pharmacological agents, etc.). (Updated on 25
CC   -MAR-2003 to correct PI field.)
XX
SQ   Sequence 1787 BP; 349 A; 603 C; 491 G; 344 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.5e-176       Length:         1787
Score:                  2694.00        Matches:        543
Percent Similarity:     97.14%         Conservative:   1
Best Local Similarity:  96.96%         Mismatches:     11
Query Match:            56.28%         Indels:         11
DB:                     2              Gaps:           2

US-09-970-944-2 (1-899) x AAV52941 (1-1787)

Qy        344 AsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspValAlaLeuTyr 363
              ||||||||||||||||||     |||||:::||||||||||||||||||||||||||||
Db          3 AACTGTACCAGTGACCTCTG-GTACACACTGCTTCTGGCCCTGAGGACGTGGCCCTCTAT 61

Qy        364 ValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuLeuValLeuIleLeuVal 383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 GTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCATCCTCGTT 121

Qy        384 TyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIleLeuThrSer 403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 TATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCA 181

Qy        404 GlyPheGlnProValSerIle-LysProSerLysAlaAspAsnProHisLeuLeuThrIl 423
              |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db        182 GGCTTCCAGCCCGTCAGCATCTAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCAT 241

Qy        423 eGlnProAspLeuSer---ThrThrThrThrTyrGlnGlySerLeuCysProArgGlnAs 442
              |||||||||||||||   ||||||||||||||||||||||||||||||||||||||||||
Db        242 CCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGA 301

Qy        442 pGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeuGlyGlyGl 462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 TGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGG 361

Qy        462 yArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheValSerArgLe 482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 CCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCT 421

Qy        482 uSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThrTyrGlyTh 502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 CTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGAC 481

Qy        502 rPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeuLeuIlePr 522
```

```
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      482 CTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGAATCAGCCTCCTCATCCC 541

Qy      522 oProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHisLysProGl 542
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      542 CCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGA 601

Qy      542 uAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleValSerCysGl 562
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      602 AGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGG 661

Qy      562 yProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGlyGluPr 582
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      662 ACCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCC 721

Qy      582 oSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySerTrpGluGl 602
            ||||||||||||||||||||    ||||||||||||||||||||||    |||||
Db      722 CAGCCCTGACAGCTGGAGCCTGGCCCTCAAAAAGCAGTCGTGCGAGGGAGC-TGGGAG-- 778

Qy      602 nAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGlnLeuGluAl 622
             |||    ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      779 -GATGT-CTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGC 836

Qy      622 aSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGluAlaLe 642
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      837 CAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCT 896

Qy      642 uSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAlaCysThrSe 662
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      897 CAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTC 956

Qy      662 rLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuLysGluVa 682
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      957 CCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGT 1016

Qy      682 lValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgValLeuHisPh 702
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1017 GGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTT 1076

Qy      702 eLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSerLeuTrpLy 722
             ||||||||||||||||||||    |||    ||||||||||||||||||||||||||||
Db     1077 -AAGGACAGTTACCACAACCTGCC-CTATC-ATCCACGATGTGCCCAGCTCCCTGTGGAA 1133

Qy      722 sSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsnGlyThrGl 742
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1134 GAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCA 1193

Qy      742 nArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSerAspLeuAl 762
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1194 GCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGC 1253

Qy      762 aCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIleAsnPheAs 782
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1254 CTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAA 1313

Qy      782 nIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAlaGlyValPr 802
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1314 CATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCC 1373

Qy      802 oAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLysIleIleSe 822
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1374 AGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTC 1433

Qy      822 rSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGlnLysLeuHi 842
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1434 CAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCA 1493

Qy      842 sLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMetIleLeuAs 862
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1494 CCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAA 1553

Qy      862 nLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAlaAlaValAl 882
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1554 CCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGC 1613

Qy      882 aGlyLeuGlyGlnProAspAlaGlyLeu-PheThrVal-SerGluAlaGluCys 899
                ||||   ||||||   |||   |||    |||||||||  ||||||||||||||
Db     1614 TGGGACTGGCCAGCAGGACGGTGGCTTCTTTCACAGTGTTCGGAGGCTGAGTGC 1667

RESULT 9
AAS75738
ID   AAS75738 standard; cDNA; 3646 BP.
XX
AC   AAS75738;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #11542.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG11551.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 11542; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 3646 BP; 930 A; 917 C; 921 G; 877 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.:              1e-174        Length:        3646
Score:                  2673.50       Matches:       507
Percent Similarity:     72.40%        Conservative:  154
Best Local Similarity:  55.53%        Mismatches:    221
```

```
Query Match:            55.85%     Indels:      31
DB:                     5          Gaps:        12

US-09-970-944-2 (1-899) x AAS75738 (1-3646)

Qy        9 ProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeuArgGlySerGlyAlaGlnGlnSer 28
Db      230 CCTGCCCTG------GCCCTGCTCAGCGCCAGCGGCACTGGCTCCGCCGCCCAAGATGAT 283

Qy       29 ---------AlaThrValAlaAsnProValProGlyAlaAsnProAspLeuLeuProHis 45
Db      284 GACTTTTTTCATGAACTCCCAGAAACTTTTCCTTCTGATCCACCTGAGCCTCTGCCACAT 343

Qy       46 PheLeuValGluProGluAspValTyrIleValLysAsnLysProValLeuLeuValCys 65
Db      344 TTCCTTATTGAGCCTGAAGAAGCTTATATTGTGAAGAATAAGCCTGTGAACCTGTACTGT 403

Qy       66 LysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGluTrpValArgGlnVal 85
Db      404 AAAGCAAGCCCTGCCACCCAGATCTATTTCAAGTGTAATAGTGAATGGGTTCATCAGAAG 463

Qy       86 AspHisValIleGluArgSerThrAspGlySerSerGlyGluProThrMetGluValArg 105
Db      464 GACCACATAGTAGATGAAAGAGTAGATGAAACTTCCGGTCTCATTGTCCGGGAAGTGAGC 523

Qy      106 IleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGluGluTyrTrpCysGln 125
Db      524 ATTGAGATTTCGCGCCAGCAAGTGGAAGAACTCTTTGGACCTGAAGATTACTGGTGCCAG 583

Qy      126 CysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAlaTyrIleArgIleAla 145
Db      584 TGTGTGGCCTGGAGCTCCGCGGGTACCACAAAGAGCCGGAAGGCGTATGTGCGCATTGCA 643

Qy      146 ArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluValSerLeuGluGlnGly 165
Db      644 TATCTACGGAAGACATTTGAGCAGGAACCCCTAGGAAAGGAAGTGTCTTTGGAACAGGAA 703

Qy      166 IleValLeuProCysArgProProGluGlyIleProProAlaGluValGluTrpLeuArg 185
Db      704 GTCTTACTCCAGTGTCGACCACCTGAAGGGATCCCAGTGGCTGAGGTGGAATGGTTGAAA 763

Qy      186 AsnGluAspLeuValAspProSerLeuAspProAsnValTyrIleThrArgGluHisSer 205
Db      764 AATGAAGACATAATTGATCCCGTTGAAGATCGGAATTTTTATATTACTATTGATCACAAC 823

Qy      206 LeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThrCysValAlaLysAsn 225
Db      824 CTCATCATAAAGCAGGCCCGACTCTCTGATACTGCAAATTACACCTGTGTTGCCAAAAAC 883

Qy      226 IleValAlaArgArgArgSerAlaSerAlaAlaValIleValTyrValAsnGlyGlyTrp 245
Db      884 ATTGTTGCCAAGAGGAAAAGTACAACTGCCACTGTCATAGTCTATGTCAACGGTGGCTGG 943

Qy      246 SerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArgGlyTrpGlnLysArg 265
Db      944 TCCACCTGGACGGAGTGGTCTGTGTGTAACAGCCGCTGTGGACGAGGGTATCAGAAACGT 1003

Qy      266 SerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPheCysGluGlyGlnAsn 285
Db     1004 ACAAGGACTTGTACCAACCCGGCACCACTCAATGGGGGTGCCTTCTGTGAAGGGCAGAGT 1063

Qy      286 ValHisAspArgThrValSerSerLeuLeuValSerValAspGlySerTrpSerProTrp 305
Db     1064 GTG---CAGAAAATAGCCTGTACTACGTTATGCCCAGTGGATGGCAGGTGGACGCCATGG 1120

Qy      306 SerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSerArgGluCysSerAsp 325
Db     1121 AGCAAGTGGTCTACTTGTGGAACTGAGTGCACCCACTGGCGCAGGAGGGAGTGCACGGCG 1180

Qy      326 ProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeuAspThrArgAsnCys 345
Db     1181 CCAGCCCCCAAGAATGGAGGCAAGGACTGCGACGGCCTCGTCTTGCAATCCAAGAACTGC 1240

Qy      346 ThrSerAspLeuCysValHisSerAlaSerGlyProGluAspValAlaLeuTyrValGly 365
Db     1241 ACTGATGGGCTTTGCATGCAGACTGCTCCTGATTCAGATGATGTTGCTCTCTATGTTGGG 1300

Qy      366 Leu---IleAlaValAlaValCysLeuValLeuLeuLeuLeuValLeuIleLeuValTyr 384
Db     1301 ATTGTGATAGCAGTGATCGTTTGCCTGGCGATCTCTGTAGTTGTGGCCTTGTTTGTGTAT 1360

Qy      385 CysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIleLeuThrSerGly 404
Db     1361 CGGAAGAATCATCGTGACTTTGAGTCAGATATTATTGACTCTTCGGCACTCAATGGGGGC 1420

Qy      405 PheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeuLeuThrIleGln 424
Db     1421 TTTCARCCTGTGAACATCAAGGCAGCAAGACAAGAT---------CTGCTGGCTGTACCC 1471

Qy      425 ProAspLeu---SerThrThrThrThrTyrGlnGlySerLeuCysProArgGlnAspGly 443
Db     1472 CCAGACCTCACGTCAGCTGCAGCCATGTACAGAGGACCTGTCTATGCCCTGCATGAC--- 1528

Qy      444 ProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeuGlyGlyGlyArg 463
Db     1529 GTCTCAGACAAAATCCCAATGACCAACTCTCCAATTCTGGATCCACTGCCCAACCTGAAA 1588

Qy      464 HisThrLeuHisHisSer-----------SerProThrSerGluAlaGluGluPheVal 479
Db     1589 ATCAAAGTGTACAACACCTCAGGTGCTGTCTCCCCCCAAGATGACCTCTCTGAGTTTACG 1648

Qy      480 SerArgLeuSer---------ThrGlnAsnTyrPhe------------------------ 488
Db     1649 TCCAAGCTGTCCCCTCAGATGACCCAGTCGTTGTTGGAGAATGAAGCCCTCAGCCTGAAG 1708

Qy      489 ---ArgSerLeuProArgGlyThr------SerAsnMetThrTyrGlyThrPheAsnPhe 505
Db     1709 AACCAGAGTCTAGCAAGGCAGACTGATCCATCCTGTACCGCATTTGGCAGCTTCAACTCG 1768

Qy      506 LeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeuLeuIleProProAspAla 525
Db     1769 CTGGGAGGTCACCTTATTGTTCCCAATTCAGGAGTCAGCTTGCTGATTCCCGCTGGGGCC 1828

Qy      526 IleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHisLysProGluAspValArg 545
Db     1829 ATTCCCCAAGGGAGAGTCTACGAAATGTATGTGACTGTACACAGGAAAGAAACTATGAGG 1888

Qy      546 LeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleValSerCysGlyProProGly 565
Db     1889 CCACCCATGGATGACTCTCAGACACTTTTGACCCCTGTGGTGAGCTGTGGGCCCCCAGGA 1948

Qy      566 ValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGlyGluProSerProAsp 585
Db     1949 GCTCTGCTCACCCGCCCCGTCGTCCTCACTATGCATCACTGCGCAGACCCCAATACCGAG 2008

Qy      586 SerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySerTrpGluGlnAspValLeu 605
Db     2009 GACTGGAAAATACTGCTCAAGAACCAGGCAGCACAGGGACAGTGGGAG---GATGTGGTG 2065

Qy      606 HisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGlnLeuGluAlaSerAlaCys 625
Db     2066 GTGGTCGGGGAGGAAAACTTCACCACCCCCTGCTACATTAAGCTGGATGCAGAGGCCTGC 2125

Qy      626 TyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGluAlaLeuSerValAla 645
Db     2126 CACATCCTCACAGAGAACCTCAGCACCTACGCCCTGGTAGGACATTCCACCACCAAAGCG 2185

Qy      646 AlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAlaCysThrSerLeuGluTyr 665
Db     2186 GCTGCAAAGCGCCTCAAGCTGGCCATCTTTGGGCCCCTGTGCTGCTCCTCGCTGGAGTAC 2245

Qy      666 AsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuLysGluValValGlnLeu 685
Db     2246 AGCATCCGAGTCTACTGTCTGGATGACACCCAGGATGCCCTGAAGGAAATTTTACATCTT 2305

Qy      686 GluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgValLeuHisPheLysAspSer 705
```

```
             |||:::|||  |||||||||||::::::||||||:::   |||||||||||   |||
Db      2306 GAGAGACAGACGGGAGGACAGCTCCTAGAAGAACCTAAGGCTCTTCATTTTAAAGGCAGC 2365

Qy       706 TyrHisAsnLeuArgLeuSerIleHisAspValProSerSerLeuTrpLysSerLysLeu 725
                |||||||||||||||||||||||||:::        ||||||||||||||||||
Db      2366 ACCCACAACCTGCGCCTGTCAATTCACGATATCGCCCATTCCCTCTGGAAGAGCAAATTG 2425

Qy       726 LeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsnGlyThrGlnArgTyrLeu 745
             |||   ||||||||||||||||||||||||:::|||:::||||:::||||||   |||
Db      2426 CTGGCTAAATATCAGGAAATTCCATTTTACCATGTTTGGAGTGGATCTCAAAGAAACCTG 2485

Qy       746 HisCysThrPheThrLeuGluArgValSerProSerThrSerAspLeuAlaCysLysLeu 765
             |||||||||||||||||||||||   |||   :::|||   :::|||   ||||||||
Db      2486 CACTGCACCTTCACTCTGGAAAGATTTAGCCTGAACACAGTGGAGCTGGTTTGCAAACTC 2545

Qy       766 TrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIleAsnPheAsnIleThrLys 785
                |||   ||||||||||||:::||||||   |||   :::|||      :::::::::
Db      2546 TGTGTGCGGCAGGTGGAAGGAGAAGGGCAGATCTTCCAGCTCAACTGCACCGTGTCAGAG 2605

Qy       786 AspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAlaGlyValProAlaLeuVal 805
             :::            :::|||     |||:::            :::      :::
Db      2606 GAACCTACTGGCATCGATTTGCCGCTGCTGGATCCTGCGAACACCATCACCACGGTCACG 2665

Qy       806 GlyProSerAlaPheLysIleProPheLeuIleArgGlnLysIleIleSerSerLeuAsp 825
             |||||||||||||||   ||||||      |||||||||||||:::   ||||||||||||
Db      2666 GGGCCCAGTGCTTTCAGCATCCCTCTCCCTATCCGGCAGAAGCTCTGTAGCAGCCTGGAT 2725

Qy       826 ProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGlnLysLeuHisLeuAspSer 845
                |||      ||||||   |||||||||   ||||||   ||||||:::||||||
Db      2726 GCCCCCCAGACGAGAGGCCATGACTGGAGGATGCTGGCCCATAAGCTGAACCTGGACAGG 2785

Qy       846 HisLeuSerPhePheAlaSerLysProSerProThrAlaMetIleLeuAsnLeuTrpGlu 865
             :::|||::::::||||||:::|||   |||||||||   :::||||||:::|||||||||
Db      2786 TACTTGAATTACTTTGCCACCAAATCCAGCCCAACTGGCGTAATCCTGGATCTTTGGGAA 2845

Qy       866 AlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAlaAlaValAlaGlyLeuGly 885
             |||::::::||||||:::||||||||||||   |||||||||   :::      :::|||
Db      2846 GCACAGAACTTCCCAGATGGAAACCTGAGCATGCTGGCAGCTGTCTTGGAAGAAATGGGA 2905

Qy       886 GlnProAspAlaGlyLeuPheThrValSerGluAlaGlu 898
             :::   :::      :::         :::|||   :::
Db      2906 AGACATGAAACGGTGGTGTCCTTAGCAGCAGAAGGGCAG 2944
```

RESULT 10
ABK92062
ID   ABK92062 standard; DNA; 2995 BP.
XX
AC   ABK92062;
XX
DT   14-AUG-2002  (first entry)
XX
DE   DNA encoding novel UNC5 receptor-like protein.
XX
KW   Cancer; metabolic disorder; skin disorder; infectious disease; anorexia;
KW   behavioral disorder; valve disease; endocrine disorder; heart disorder;
KW   blood disorder; anxiety disorder; brain disorder; inflammatory disorder;
KW   neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW   immune disorder; haematopoietic disorder; dyslipidaemia;
KW   metabolic disturbance; obesity; metabolic syndrome X; wasting disorder;
KW   gene therapy; transgenic animal; human; gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO200229058-A2.
XX
PD   11-APR-2002.
XX
PF   05-OCT-2001; 2001WO-US031248.
XX
PR   05-OCT-2000; 2000US-0238323P.
PR   05-OCT-2000; 2000US-0238325P.
PR   06-OCT-2000; 2000US-0238372P.
PR   06-OCT-2000; 2000US-0238373P.
PR   06-OCT-2000; 2000US-0238379P.
PR   06-OCT-2000; 2000US-0238382P.
PR   06-OCT-2000; 2000US-0238383P.
PR   06-OCT-2000; 2000US-0238384P.
PR   06-OCT-2000; 2000US-0238397P.
PR   06-OCT-2000; 2000US-0238400P.
PR   06-OCT-2000; 2000US-0238401P.
PR   06-OCT-2000; 2000US-0238402P.
PR   14-MAR-2001; 2001US-0275892P.
PR   08-JUN-2001; 2001US-0296860P.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA, Taupier RJ, Burgess CE, Zerhusen BD, Mezes PS;
PI   Rastelli L, Malyankar UM, Grosse WM, Alsobrook JP, Lepley DM;
PI   Spytek KA, Li L, Edinger S, Gerlach V, Ellerman K, Macdougall J;
PI   Gunther E, Millet I, Stone D, Smithson G, Szekeres ES;
XX
DR   WPI; 2002-444103/47.
DR   P-PSDB; ABG61795.
XX
PT   Novel isolated polypeptide, designated NOVX, useful for treating or
PT   preventing cancer, diabetes, obesity, dyslipidemia, anorexia, and
PT   metabolic, neurodegenerative, immune and hematopoietic disorders.
XX
PS   Claim 9; Page 93; 316pp; English.
XX
CC   The invention describes an isolated polypeptide (I), useful in the
CC   manufacture of a medicament for treating a syndrome associated with a
CC   human disease. (I), the polynucleotide encoding it (II) and an antibody
CC   (III) to (I) are useful for treating or preventing cancer, metabolic
CC   disorders, skin disorders, infectious disease, anorexia, behavioral
CC   disorders, valve diseases, endocrine disorders, heart and blood
CC   disorders, anxiety disorders, brain disorders, inflammatory disorders,
CC   neurodegenerative disorder, Alzheimer's disease, Parkinson's disorder,
CC   immune disorders, haematopoietic disorders, and the various
CC   dyslipidaemias, metabolic disturbances associated with obesity, and the
CC   metabolic syndrome X and wasting disorders associated with chronic
CC   diseases. (I), (II) or (III) are useful in screening assays, detection
CC   assays (e.g., chromosomal mapping, tissue typing, forensic biology),
CC   predictive medicine (e.g., diagnostic assays and prognostic assays), and
CC   in methods of treatment (e.g., therapeutic and prophylactic). (I) is
CC   useful as immunogen to produce antibodies immunospecific for (I), to
CC   screen for potential agonist and antagonist compounds, and as bait
CC   protein in a two-hybrid or three-hybrid assay. (II) is useful in gene
CC   therapy, to express (I), to detect (I) mRNA or a genetic lesion in a (I)
CC   gene, and to modulate activity of (I). A cell containing a vector
CC   exressing (I) is useful for producing non-human transgenic animals. This
CC   sequence encodes a novel human polypeptide described in the invention
XX
SQ   Sequence 2995 BP; 571 A; 984 C; 925 G; 515 T; 0 U; 0 Other;

```
Alignment Scores:
Pred. No.:              3.38e-162     Length:        2995
Score:                  2490.50       Matches:       496
Percent Similarity:     67.86%        Conservative:  150
Best Local Similarity:  52.10%        Mismatches:    247
Query Match:            52.03%        Indels:        59
DB:                     6             Gaps:          18
```

US-09-970-944-2 (1-899) x ABK92062 (1-2995)

```
Qy         1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrp--- 19
             |||      |||   |||         |||||||||      :::|||      |||
Db        64 ATGGGGGCCCGGAGCGGAGCTCGGGGCGCGCTGCTGCTGGCACTGCTGCTCTGCTGGGAC 123

Qy        20 ------LeuArgGlySerGlyAlaGlnGlnSerAlaThrVal---AlaAsnProValPro 36
                   |||      :::|||   ::::::|||      |||      :::      |||
Db       124 CCGAGGCTGAGCCAAGCAGGTAGGAAGCGATCGGGTGAGGTGCTCCCTGACTCCTTCCCG 183

Qy        37 GlyAlaAsnProAspLeuLeuProHisPheLeuValGluProGluAspValTyrIleVal 56
                |||      :::   ||||||:::||||||   ||||||:::|||   ||||||||
Db       184 TCAGCGCCAGCAGAGCCGCTGCCCTACTTCCTGCAGGAGCCACAGGACGCCTACATTGTG 243
```

```
Qy     57 LysAsnLysProValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLys 76
Db    244 AAGAACAAGCCTGTGGAGCTCCGCTGCCGCGCCTTCCCCGCCACACAGATCTACTTCAAG 303

Qy     77 CysAsnGlyGluTrpValArgGlnValAspHisValIleGluArgSerThrAspGlySer 96
Db    304 TGCAACGGCGAGTGGGTCAGCCAGAACGACCACGTCACACAGGAAGGCCTGGATGAGGCC 363

Qy     97 Ser---------------GlyGluProThrMetGluValArgIleAsnValSerArgGln 111
Db    364 ACCCTGGGGGCGCGGGGCGGCCTGCGGGTGCGCGAGGTGCAGATCGAGGTGTCGCGGCAG 423

Qy    112 GlnValGluLysValPheGlyLeuGluGluTyrTrpCysGlnCysValAlaTrpSerSer 131
Db    424 CAGGTGGAGGAGCTCTTTGGGCTGGAGGATTACTGGTGCCAGTGCGTGGCCTGGAGCTCC 483

Qy    132 SerGlyThrThrLysSerGlnLysAlaTyrIleArgIleAlaArgLeuArgLysAsnPhe 151
Db    484 GCGGGCACCACCAAGAGTCGCCGAGCCTACGTCCGCATCGCCTGTCTGCGCAAGAACTTC 543

Qy    152 GluGlnGluProLeuAlaLysGluValSerLeuGluGlnGlyIleValLeuProCysArg 171
Db    544 GATCAGGAGCCTCTGGGCAAGGAGGTGCCCCTGGACCATGAGGTTCTCCTGCAGTGCCGC 603

Qy    172 ProProGluGlyIleProProAlaGluValGluTrpLeuArgAsnGluAspLeuValAsp 191
Db    604 CCGCCGGAGGGGGTGCCTGTGGCCGAGGTGGAATGGCTCAAGAATGAGGATGTCATCGAC 663

Qy    192 ProSerLeuAspProAsnValTyrIleThrArgGluHisSerLeuValValArgGlnAla 211
Db    664 CCCACCCAGGACACCAACTTCCTGCTCACCATCGACCACAACCTCATCATCCGCCAGGCC 723

Qy    212 ArgLeuAlaAspThrAlaAsnTyrThrCysValAlaLysAsnIleValAlaArgArgArg 231
Db    724 CGCCTGTCGGACACTGCCAACTATACCTGCGTGGCCAAGAACATCGTGGCCAAACGCCGG 783

Qy    232 SerAlaSerAlaAlaValIleValTyrValAsnGlyGlyTrpSerThrTrpThrGluTrp 251
Db    784 AGCACCACTGCCACCGTCATCGTCTACGTGAATGGCGGCTGGTCCAGCTGGGCAGAGTGG 843

Qy    252 SerValCysSerAlaSerCysGlyArgGlyTrpGlnLysArgSerArgSerCysThrAsn 271
Db    844 TCACCCTGCTCCAACCGCTGTGGCCGAGGCTGGCAGAAGCGCACCCGGACCTGCACCAAC 903

Qy    272 ProAlaProLeuAsnGlyGlyAlaPheCysGluGlyGlnAsnValHisAspArgThrVal 291
Db    904 CCCGCTCCACTCAACGGAGGGGCCTTCTGCGAGGGCCAG---GCATTCCAGAAGACCGCC 960

Qy    292 SerSerLeuLeuValSerValAspGlySerTrpSerProTrpSerLysTrpSerAlaCys 311
Db    961 TGCACCACCATCTGCCCAGTCGATGGGGCGTGGACGGAGTGGAGCAAGTGGTCAGCCTGC 1020

Qy    312 GlyLeuAspCysThrHisTrpArgSerArgGluCysSerAspProAlaProArgAsnGly 331
Db   1021 AGCACTGAGTGTGCCCACTGGCGTAGCCGCGAGTGCATGGCGCCCCCACCCCAGAACGGA 1080

Qy    332 GlyGluGluCysGlnGlyThrAspLeuAspThrArgAsnCysThrSerAspLeuCysVal 351
Db   1081 GGCCGTGACTGCAGCGGGACGCTGCTCGACTCTAAGAACTGCACAGATGGGCTGTGCATG 1140

Qy    352 HisSer-----------------------AlaSerGlyProGluAspValAlaLeuTyr 363
Db   1141 CAAAGTGAGCCTGTCCCCGCAGTGCTGGAGGCCTCAGGG------GATGCGGCGCTGTAT 1194

Qy    364 ValGlyLeu---IleAlaValAlaValCysLeuValLeuLeuLeuLeuValLeuIleLeu 382
Db   1195 GCGGGGCTCGTGGTGGCCATCTTCGTGGTCGTGGCAATCCTCATGGCGGTGGGGGTGGTG 1254

Qy    383 ValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSer---IleLeu 401
Db   1255 GTGTACCGCCGCAACTGCCGTGACTTCGACACAGACATCACTGACTCATCTGCTGCCCTG 1314

Qy    402 ThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeuLeu 421
Db   1315 ACTGGTGGTTTCCACCCCGTCAACTTTAAGACGGCAAGGCCCAGTAACCCGCAGCTCCTA 1374

Qy    422 ------ThrIleGlnProAspLeuSerThrThrThr---ThrTyrGlnGlySerLeuCys 438
Db   1375 CACCCCTCTGTGCCTCCTGACCTGACAGCCAGCGCCGGCATCTACCGCGGACCCGTGTAT 1434

Qy    439 ProArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerPro 458
Db   1435 GCCCTGCAGGACTCC---ACCGACAAAATCCCCATGACCAACTCTCCTCTGCTGGACCCC 1491

Qy    459 LeuGlyGlyGlyArgHisThrLeuHisHisSerSerProThr----------------- 472
Db   1492 TTACCCAGCCTTAAGGTCAAGGTCTACAGCTCCAGCACCACGGGCTCTGGGCCAGGCCTG 1551

Qy    473 SerGluAlaGluGluPheValSerArgLeuSerThrGlnAsnTyr-------------- 487
Db   1552 GCAGATGGGGCTGACCTGCTGGGGGTCTTGCCGCCTGGCACATACCCTAGCGATTTCGCC 1611

Qy    488 --------------------PheArgSer------------------------------ 490
Db   1612 CGGGACACCCACTTCCTGCACCTGCGCAGCGCCAGCCTCGGTTCCCAGCAGCTCTTGGGC 1671

Qy    491 LeuProArgGlyThrSerAsnMetThrTyrGlyThrPheAsnPheLeuGlyGlyArgLeu 510
Db   1672 CTGCCCCGAGACCCAGGGAGCAGCGTCAGCGGCACCTTTGGCTGCCTGGGTGGGAGGCTC 1731

Qy    511 MetIleProAsnThrGlyIleSerLeuLeuIleProProAspAlaIleProArgGlyLys 530
Db   1732 AGCATCCCCGGCACAGGTGTCAGCTTGCTGGTGCCCAATGGAGCCATTCCCCAGGGCAAG 1791

Qy    531 IleTyrGluIleTyrLeuThrLeuHisLysProGluAspValArgLeuProLeuAla--- 549
Db   1792 TTCTACGAGATGTATCTACTCATCAACAAGGCAGAAAGTACC---CTGCCGCTTTCAGAA 1848

Qy    550 GlyCysGlnThrLeuLeuSerProIleValSerCysGlyProProGlyValLeuLeuThr 569
Db   1849 GGGACCCAGACAGTATTGAGCCCCTCGGTGACCTGTGGACCCACAGGCCTCCTGCTGTGC 1908

Qy    570 ArgProValIleLeuAlaMetAspHisCysGlyGluProSerProAspSerTrpSerLeu 589
Db   1909 CGCCCCGTCATCCTCACCATGCCCCACTGTGCCGAAGTCAGTGCCCGTGACTGGATCTTT 1968

Qy    590 ArgLeuLysLysGlnSerCysGluGlySerTrpGluGlnAspValLeuHisLeuGlyGlu 609
Db   1969 CAGCTCAAGACCCAGGCCCACCAGGGCCACTGGGAGCAGGAGGTGGTGACCCTGGATGAG 2028

Qy    610 GluAlaProSerHisLeuTyrTyrCysGlnLeuGluAlaSerAlaCysTyrValPheThr 629
Db   2029 GAGACCCTGAACACACCCTGCTACTGCCAGCTGGAGCCCAGGGCCTGTCACATCCTGCTG 2088

Qy    630 GluGlnLeuGlyArgPheAlaLeuValGlyGluAlaLeuSerValAlaAlaAlaLysArg 649
Db   2089 GACCAGCTGGGCACCTACGTGTTCACGGGCGAGTCCTATTCCCGCTCAGCAGTCAAGCGG 2148

Qy    650 LeuLysLeuLeuLeuPheAlaProValAlaCysThrSerLeuGluTyrAsnIleArgVal 669
Db   2149 CTCCAGCTGGCCGTCTTCGCCCCCGCCCTCTGCACCTCCCTGGAGTACAGCCTCCGGGTC 2208

Qy    670 TyrCysLeuHisAspThrHisAspAlaLeuLysGluValValGlnLeuGluLysGlnLeu 689
Db   2209 TACTGCCTGGAGGACACGCCTGTAGCACTGAAGGAGGTGCTGGAGCTGGAGCGGACTCTG 2268

Qy    690 GlyGlyGlnLeuIleGlnGluProArgValLeuHisPheLysAspSerTyrHisAsnLeu 709
Db   2269 GGCGGATACTTGGTGGAGGAGCCGAAACCGCTAATGTTCAAGGACAGTTACCACAACCTG 2328

Qy    710 ArgLeuSerIleHisAspValProSerSerLeuTrpLysSerLysLeuLeuValSerTyr 729
Db   2329 CGCCTCTCCCTCCATGACCTCCCCCATGCCCATTGGAGGAGCAAGCTGCTGGCCAAATAC 2388

Qy    730 GlnGluIleProPheTyrHisIleTrpAsnGlyThrGlnArgTyrLeuHisCysThrPhe 749
```

```
Db     2389 CAGGAGATCCCCTTCTATCACATTTGGAGTGGCAGCCAGAAGGCCCTCCACTGCACTTTC 2448

Qy      750 ThrLeuGluArgValSerProSerThrSerAspLeuAlaCysLysLeuTrpValTrpGln 769
Db     2449 ACCCTGGAGAGGCACAGCTTGGCCTCCACAGAGCTCACCTGCAAGATCTGCGTGCGGCAA 2508

Qy      770 ValGluGlyAspGlyGlnSerPheSerIleAsnPheAsnIleThrLysAspThrArgPhe 789
Db     2509 GTGGAAGGGGAGGGCCAGATATTCCAGCTGCATACCACTCTGGCA---GAGACACCTGCT 2565

Qy      790 AlaGluLeuLeuAlaLeuGluSerGluAlaGly------ValProAlaLeuValGlyPro 807
Db     2566 GGCTCCCTGGACACTCTCTGCTCTGCCCCTGGCAGCACTGTCACCACCCAGCTGGGACCT 2625

Qy      808 SerAlaPheLysIleProPheLeuIleArgGlnLysIleIleSerSerLeuAspProPro 827
Db     2626 TATGCCTTCAAGATCCCACTGTCCATCCGCCAGAAGATATGCAACAGCCTAGATGCCCCC 2685

Qy      828 CysArgArgGlyAlaAspTrpArgThrLeuAlaGlnLysLeuHisLeuAspSerHisLeu 847
Db     2686 AACTCACGGGGCAATGACTGGCGGATGTTAGCACAGAAGCTCTCTATGGACCGGTACCTG 2745

Qy      848 SerPhePheAlaSerLysProSerProThrAlaMetIleLeuAsnLeuTrpGluAlaArg 867
Db     2746 AATTACTTTGCCACCAAAGCGAGCCCCACGGGTGTGATCCTGGACCTCTGGGAAGCTCTG 2805

Qy      868 HisPheProAsnGlyAsnLeuSerGlnLeuAlaAlaAlaValAlaGlyLeuGlyGlnPro 887
Db     2806 CAGCAGGACGATGGGGACCTCAACAGCCTGGCGAGTGCCTTGGAGGAGATGGGCAAGAGT 2865

Qy      888 AspAlaGlyLeuPheThrValSerGluAlaGluCys 899
Db     2866 GAGATGCTGGTGGCTGTGGCCACCGACGGGGACTGC 2901
```

RESULT 11
ABT06279

```
ID   ABT06279 standard; cDNA; 2860 BP.
XX
AC   ABT06279;
XX
DT   24-OCT-2002  (first entry)
XX
DE   Human NOV1a coding sequence.
XX
KW   Human; NOVX; autoimmune disease; cancer; infection; inflammatory disease;
KW   storage disorder; muscle disorder; neurodegenerative disorder; nootropic;
KW   developmental defect; neuroprotective; antiparkinsonian; hypotensive;
KW   hypertensive; haemostatic; cardiant; antianginal; dermatological;
KW   immunosuppressive; antiinflammatory; virucide; antibacterial; anti-HIV;
KW   antiparasitic; antiallergic; antiasthmatic; antirheumatic; antiarthritic;
KW   vulnerary; anorectic; antidiabetic; immunomodulator; antipsoriatic;
KW   nephrotropic; kerolytic; antiulcer; cerebroprotective; anticonvulsant;
KW   antiinfertility; antimanic; antidepressant; metabolic; cytostatic;
KW   tranquilizer; analgesic; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200257450-A2.
XX
PD   25-JUL-2002.
XX
PF   29-NOV-2001; 2001WO-US048922.
XX
PR   29-NOV-2000; 2000US-0253834P.
PR   30-NOV-2000; 2000US-0250926P.
PR   25-JAN-2001; 2001US-0264180P.
PR   20-AUG-2001; 2001US-0313656P.
PR   05-OCT-2001; 2001US-0327456P.
PR   28-NOV-2001; 2001US-00327456.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Edinger S,   Macdougall JR,   Millet I,   Ellerman K,   Stone DJ;
PI   Gerlach V,   Grosse WM,   Alsobrook JP,   Lepley DM,   Rieger D,   Burgess CE;
PI   Casman SJ,   Spytek KA,   Boldog FL,   Li L,   Padigaru M,   Mishra V;
PI   Patturajan M,   Shenoy S,   Rastelli L,   Tchernev VT,   Vernet CAM;
PI   Zerhusen BD,   Malyankar UM,   Guo X,   Miller CE,   Gangolli EA;
XX
DR   WPI; 2002-590741/63.
DR   P-PSDB; AAO18734.
XX
PT   Novel isolated polypeptide, designated NOVX, useful for treating or
PT   preventing in NOVX-associated disorders e.g. cardiomyopathy,
PT   atherosclerosis, diabetes, cancer, allergy, asthma, Crohn's disease.
XX
PS   Claim 9; Page 11; 353pp; English.
XX
CC   The present invention provides the protein and coding sequences of
CC   several novel human proteins, designated NOVX. These can be used in the
CC   treatment of, amongst others, cancers, autoimmune diseases, infections,
CC   inflammatory diseases, storage disorders, muscle disorders,
CC   neurodegenerative diseases and developmental defects. The present
CC   sequence is a coding sequence of the invention
XX
SQ   Sequence 2860 BP; 555 A; 950 C; 860 G; 495 T; 0 U; 0 Other;
```

Alignment Scores:

| | | | |
|---|---|---|---|
| Pred. No.: | 8.93e-162 | Length: | 2860 |
| Score: | 2484.00 | Matches: | 493 |
| Percent Similarity: | 68.30% | Conservative: | 149 |
| Best Local Similarity: | 52.45% | Mismatches: | 250 |
| Query Match: | 51.89% | Indels: | 48 |
| DB: | 6 | Gaps: | 18 |

US-09-970-944-2 (1-899) x ABT06279 (1-2860)

```
Qy        1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrp--- 19
Db       59 ATGGGGGCCCGGAGCGGAGCTCGGGGCGCGCTGCTGCTGGCACTGCTGCTCTGCTGGGAC 118

Qy       20 ------LeuArgGlySerGlyAlaGlnGln---SerAlaThrValAlaAsnProValPro 36
Db      119 CCGAGGCTGAGCCAAGCAGGCACTGATTCTGGCAGCGAGGTGCTCCCTGACTCCTTCCCG 178

Qy       37 GlyAlaAsnProAspLeuLeuProHisPheLeuValGluProGluAspValTyrIleVal 56
Db      179 TCAGCGCCAGCAGAGCCGCTGCCCTACTTCCTGCAGGAGCCACAGGACGCCTACATTGTG 238

Qy       57 LysAsnLysProValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLys 76
Db      239 AAGAACAAGCCTGTGGAGCTCCGCTGCCGCGCCTTCCCCGCCACACAGATCTACTTCAAG 298

Qy       77 CysAsnGlyGluTrpValArgGlnValAspHisValIleGluArgSerThrAspGlySer 96
Db      299 TGCAACGGCGAGTGGGTCAGCCAGAACGACCACGTCACACAGGAAGGCCTGGATGAGGCC 358

Qy       97 SerGlyGluProThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysVal 116
Db      359 ACCGGTCTGCGGGTGCGCGAGGTGCAGATCGAGGTGTCGCGGCAGCAGGTGGAGGAGCTC 418

Qy      117 PheGlyLeuGluGluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLys 136
Db      419 TTTGGGCTGGAGGATTACTGGTGCCAGTGCGTGGCCTGGAGCTCCGCGGGCACCACCAAG 478

Qy      137 SerGlnLysAlaTyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeu 156
Db      479 AGTCGCCGAGCCTACGTCCGCATCGCCTACCTGCGCAAGAACTTCGATCAGGAGCCTCTG 538

Qy      157 AlaLysGluValSerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIle 176
Db      539 GGCAAGGAGGTGCCCCTGGACCATGAGGTTCTCCTGCAGTGCCGCCCGCCGGAGGGGGTG 598

Qy      177 ProProAlaGluValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspPro 196
Db      599 CCTGTGGCCGAGGTGGAATGGCTCAAGAATGAGGATGTCATCGACCCCACCCAGGACACC 658
```

```
Qy    197 AsnValTyrIleThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThr 216
Db    659 AACTTCCTGCTCACCATCGACCACAACCTCATCATCCGCCAGGCCCGCCTGTCGGACACT 718

Qy    217 AlaAsnTyrThrCysValAlaLysAsnIleValAlaArgArgArgSerAlaSerAlaAla 236
Db    719 GCCAACTATACCTGCGTGGCCAAGAACATCGTGGCCAAACGCCGGAGCACCACTGCCACC 778

Qy    237 ValIleValTyrValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAla 256
Db    779 GTCATCGTCTACGTGAATGGCGGCTGGTCCAGCTGGGCAGAGTGGTCACCCTGCTCCAAC 838

Qy    257 SerCysGlyArgGlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsn 276
Db    839 CGCTGTGGCCGAGGCTGGCAGAAGCGCACCCGGACCTGCACCAACCCCGCTCCACTCAAC 898

Qy    277 GlyGlyAlaPheCysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuVal 296
Db    899 GGAGGGGCCTTCTGCGAGGGCCAG---GCATTCCAGAAGACCGCCTGCACCACCATCTGC 955

Qy    297 SerValAspGlySerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThr 316
Db    956 CCAGTCGATGGGGCGTGGACGGAGTGGAGCAAGTGGTCAGCCTGCAGCACTGAGTGTGCC 1015

Qy    317 HisTrpArgSerArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGln 336
Db   1016 CACTGGCGTAGCCGCGAGTGCATGGCGCCCCCACCCCAGAACGGAGGCCGTGACTGCAGC 1075

Qy    337 GlyThrAspLeuAspThrArgAsnCysThrSerAspLeuCysValHis---SerAlaSer 355
Db   1076 GGGACGCTGCTCGACTCTAAGAACTGCACAGATGGGCTGTGCATGCAACTGGAGGCCTCA 1135

Qy    356 GlyProGluAspValAlaLeuTyrValGlyLeu---IleAlaValAlaValCysLeuVal 374
Db   1136 GGG------GATGCGGCGCTGTATGCGGGGCTCGTGGTGGCCATCTTCGTGGTCGTGGCA 1189

Qy    375 LeuLeuLeuLeuValLeuIleLeuValTyrCysArgLysLysGluGlyLeuAspSerAsp 394
Db   1190 ATCCTCATGGCGGTGGGGGTGGTGGTGTACCGCCGCAACTGCCGTGACTTCGACACAGAC 1249

Qy    395 ValAlaAspSerSer---IleLeuThrSerGlyPheGlnProValSerIleLysProSer 413
Db   1250 ATCACTGACTCATCTGCTGCCCTGACTGGTGGTTTCCACCCCGTCAACTTTAAGACGGCA 1309

Qy    414 LysAlaAspAsnProHisLeuLeu------ThrIleGlnProAspLeuSerThrThrThr 431
Db   1310 AGGCCCAGTAACCCGCAGCTCCTACACCCCTCTGTGCCTCCTGACCTGACAGCCAGCGCC 1369

Qy    432 ---ThrTyrGlnGlySerLeuCysProArgGlnAspGlyProSerProLysPheGlnLeu 450
Db   1370 GGCATCTACCGCGGACCCGTGTATGCCCTGCAGGACTCC---ACCGACAAAATCCCCATG 1426

Qy    451 ThrAsnGlyHisLeuLeuSerProLeuGlyGlyGlyArgHisThrLeuHisHisSerSer 470
Db   1427 ACCAACTCTCCTCTGCTGGACCCCTTACCCAGCCTTAAGGTCAAGGTCTACAGCTCCAGC 1486

Qy    471 ProThr------------------SerGluAlaGluGluPheValSerArgLeuSerThr 484
Db   1487 ACCACGGGCTCTGGGCCAGGCCTGGCAGATGGGGCTGACCTGCTGGGGGTCTTGCCGCCT 1546

Qy    485 GlnAsnTyr---------------------------------PheArgSer------ 490
Db   1547 GGCACATACCCTAGCGATTTCGCCCGGGACACCCACTTCCTGCACCTGCGCAGCGCCAGC 1606

Qy    491 -----------------------LeuProArgGlyThrSerAsnMetThrTyrGlyThr 502
Db   1607 CTCGGTTCCCAGCAGCTCTTGGGCCTGCCCCGAGACCCAGGGAGCAGCGTCAGCGGCACC 1666

Qy    503 PheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeuLeuIlePro 522
Db   1667 TTTGGCTGCCTGGGTGGGAGGCTCAGCATCCCCGGCACAGGGGTCAGCTTGCTGGTGCCC 1726

Qy    523 ProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHisLysProGlu 542
Db   1727 AATGGAGCCATTCCCCAGGGCAAGTTCTACGAGATGTATCTACTCATCAACAAGGCAGAA 1786

Qy    543 AspValArgLeuProLeuAla---GlyCysGlnThrLeuLeuSerProIleValSerCys 561
Db   1787 AGTACC---CTGCCGCTTTCAGAAGGGACCCAGACAGTATTGAGCCCCTCGGTGACCTGT 1843

Qy    562 GlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGlyGlu 581
Db   1844 GGACCCACAGGCCTCCTGCTGTGCCGCCCCGTCATCCTCACCATGCCCCACTGTGCCGAA 1903

Qy    582 ProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySerTrpGlu 601
Db   1904 GTCAGTGCCCGTGACTGGATCTTTCAGCTCAAGACCCAGGCCCACCAGGGCCACTGGGAG 1963

Qy    602 GlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGlnLeuGlu 621
Db   1964 ---GAGGTGGTGACCCTGGATGAGGAGACCCTGAACACACCCTGCTACTGCCAGCTGGAG 2020

Qy    622 AlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGluAla 641
Db   2021 CCCAGGGCCTGTCACATCCTGCTGGACCAGCTGGGCACCTACGTGTTCACGGGCGAGTCC 2080

Qy    642 LeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAlaCysThr 661
Db   2081 TATTCCCGCTCAGCAGTCAAGCGGCTCCAGCTGGCCGTCTTCGCCCCCGCCCTCTGCACC 2140

Qy    662 SerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuLysGlu 681
Db   2141 TCCCTGGAGTACAGCCTCCGGGTCTACTGCCTGGAGGACACGCCTGTAGCACTGAAGGAG 2200

Qy    682 ValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgValLeuHis 701
Db   2201 GTGCTGGAGCTGGAGCGGACTCTGGGCGGATACTTGGTGGAGGAGCCGAAACCGCTAATG 2260

Qy    702 PheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSerLeuTrp 721
Db   2261 TTCAAGGACAGTTACCACAACCTGCGCCTCTCCCTCCATGACCTCCCCCATGCCCATTGG 2320

Qy    722 LysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsnGlyThr 741
Db   2321 AGGAGCAAGCTGCTGGCCAAATACCAGGAGATCCCCTTCTATCACATTTGGAGTGGCAGC 2380

Qy    742 GlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSerAspLeu 761
Db   2381 CAGAAGGCCCTCCACTGCACTTTCACCCTGGAGAGGCACAGCTTGGCCTCCACAGAGCTC 2440

Qy    762 AlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIleAsnPhe 781
Db   2441 ACCTGCAAGATCTGCGTGCGGCAAGTGGAAGGGGAGGGCCAGATATTCCAGCTGCATACC 2500

Qy    782 AsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAlaGly--- 800
Db   2501 ACTCTGGCA---GAGACACCTGCTGGCTCCCTGGACACTCTCTGCTCTGCCCCTGGCAGC 2557

Qy    801 ---ValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
Db   2558 ACTGTCACCACCCAGCTGGGACCTTATGCCTTCAAGATCCCACTGTCCATCCGCCAGAAG 2617

Qy    820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839
Db   2618 ATATGCAACAGCCTAGATGCCCCCAACTCACGGGGCAATGACTGGCGGATGTTAGCACAG 2677

Qy    840 LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859
Db   2678 AAGCTCTCTATGGACCGGTACCTGAATTACTTTGCCACCAAAGCGAGCCCCACGGGTGTG 2737

Qy    860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879
Db   2738 ATCCTGGACCTCTGGGAAGCTCTGCAGCAGGACGATGGGGACCTCAACAGCCTGGCGAGT 2797

Qy    880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
```

```
              |||:::       :::|||:::    :::       :::          ::::::   :::|||
Db       2798 GCCTTGGAGGAGATGGGCAAGAGTGAGATGCTGGTGGCTGTGGCCACCGACGGGGACTGC 2857

RESULT 12
ABT06280
ID   ABT06280 standard; cDNA; 2860 BP.
XX
AC   ABT06280;
XX
DT   24-OCT-2002  (first entry)
XX
DE   Human NOV1b coding sequence.
XX
KW   Human; NOVX; autoimmune disease; cancer; infection; inflammatory disease;
KW   storage disorder; muscle disorder; neurodegenerative disorder; nootropic;
KW   developmental defect; neuroprotective; antiparkinsonian; hypotensive;
KW   hypertensive; haemostatic; cardiant; antianginal; dermatological;
KW   immunosuppressive; antiinflammatory; virucide; antibacterial; anti-HIV;
KW   antiparasitic; antiallergic; antiasthmatic; antirheumatic; antiarthritic;
KW   vulnerary; anorectic; antidiabetic; immunomodulator; antipsoriatic;
KW   nephrotropic; kerolytic; antiulcer; cerebroprotective; anticonvulsant;
KW   antiinfertility; antimanic; antidepressant; metabolic; cytostatic;
KW   tranquilizer; analgesic; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200257450-A2.
XX
PD   25-JUL-2002.
XX
PF   29-NOV-2001; 2001WO-US048922.
XX
PR   29-NOV-2000; 2000US-0253834P.
PR   30-NOV-2000; 2000US-0250926P.
PR   25-JAN-2001; 2001US-0264180P.
PR   20-AUG-2001; 2001US-0313656P.
PR   05-OCT-2001; 2001US-0327456P.
PR   28-NOV-2001; 2001US-00327456.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Edinger S,  Macdougall JR,  Millet I,  Ellerman K,  Stone DJ;
PI   Gerlach V,  Grosse WM,  Alsobrook JP,  Lepley DM,  Rieger D,  Burgess CE;
PI   Casman SJ,  Spytek KA,  Boldog FL,  Li L,  Padigaru M,  Mishra V;
PI   Patturajan M,  Shenoy S,  Rastelli L,  Tchernev VT,  Vernet CAM;
PI   Zerhusen BD,  Malyankar UM,  Guo X,  Miller CE,  Gangolli EA;
XX
DR   WPI; 2002-590741/63.
DR   P-PSDB; AAO18735.
XX
PT   Novel isolated polypeptide, designated NOVX, useful for treating or
PT   preventing in NOVX-associated disorders e.g. cardiomyopathy,
PT   atherosclerosis, diabetes, cancer, allergy, asthma, Crohn's disease.
XX
PS   Claim 9; Page 14; 353pp; English.
XX
CC   The present invention provides the protein and coding sequences of
CC   several novel human proteins, designated NOVX. These can be used in the
CC   treatment of, amongst others, cancers, autoimmune diseases, infections,
CC   inflammatory diseases, storage disorders, muscle disorders,
CC   neurodegenerative diseases and developmental defects. The present
CC   sequence is a coding sequence of the invention
XX
SQ   Sequence 2860 BP; 557 A; 950 C; 858 G; 495 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              8.93e-162      Length:         2860
Score:                  2484.00        Matches:        493
Percent Similarity:     68.30%         Conservative:   149
Best Local Similarity:  52.45%         Mismatches:     250
Query Match:            51.89%         Indels:         48
DB:                     6              Gaps:           18
```

```
US-09-970-944-2 (1-899) x ABT06280 (1-2860)

Qy          1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrp--- 19
Db         59 ATGGGGGCCCGGAGCGGAGCTCGGGGCGCGCTGCTGCTGGCACTGCTGCTCTGCTGGGAC 118

Qy         20 ------LeuArgGlySerGlyAlaGlnGln---SerAlaThrValAlaAsnProValPro 36
Db        119 CCGAGGCTGAGCCAAGCAGGCACTGATTCTGGCAGCGAGGTGCTCCCTGACTCCTTCCCG 178

Qy         37 GlyAlaAsnProAspLeuLeuProHisPheLeuValGluProGluAspValTyrIleVal 56
Db        179 TCAGCGCCAGCAGAGCCGCTGCCCTACTTCCTGCAGGAGCCACAGGACGCCTACATTGTG 238

Qy         57 LysAsnLysProValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLys 76
Db        239 AAGAACAAGCCTGTGGAGCTTCGCTGCCGCGCCTTCCCCGCCACACAGATCTACTTCAAG 298

Qy         77 CysAsnGlyGluTrpValArgGlnValAspHisValIleGluArgSerThrAspGlySer 96
Db        299 TGCAACGGCGAGTGGGTCAGCCAGAACGACCACGTCACACAGGAAGGCCTGGATGAGGCC 358

Qy         97 SerGlyGluProThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysVal 116
Db        359 ACCGGCCTGCGGGTGCGCGAGGTGCAGATCGAGGTGTCGCGGCAGCAGGTGGAGGAGCTC 418

Qy        117 PheGlyLeuGluGluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLys 136
Db        419 TTTGGGCTGGAGGATTACTGGTGCCAGTGCGTGGCCTGGAGCTCCGCAGGCACCACCAAG 478

Qy        137 SerGlnLysAlaTyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeu 156
Db        479 AGTCGCCGAGCCTACGTCCGCATCGCCTACCTGCGCAAGAACTTCGATCAGGAGCCTCTG 538

Qy        157 AlaLysGluValSerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIle 176
Db        539 GGCAAGGAGGTGCCCCTGGACCATGAGGTTCTCCTGCAGTGCCGCCCGCCGGAGGGGGTG 598

Qy        177 ProProAlaGluValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspPro 196
Db        599 CCTGTGGCCGAGGTGGAATGGCTCAAGAATGAGGATGTCATCGACCCCACCCAGGACACC 658

Qy        197 AsnValTyrIleThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThr 216
Db        659 AACTTCCTGCTCACCATCGACCACAACCTCATCATCCGCCAGGCCCGCCTGTCGGACACT 718

Qy        217 AlaAsnTyrThrCysValAlaLysAsnIleValAlaArgArgArgSerAlaSerAlaAla 236
Db        719 GCCAACTATACCTGCGTGGCCAAGAACATCGTGGCCAAACGCCGGAGCACCACTGCCACC 778

Qy        237 ValIleValTyrValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAla 256
Db        779 GTCATCGTCTACGTGAATGGCGGCTGGTCCAGCTGGGCAGAGTGGTCACCCTGCTCCAAC 838

Qy        257 SerCysGlyArgGlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsn 276
Db        839 CGCTGTGGCCGAGGCTGGCAGAAGCGCACCCGGACCTGCACCAACCCCGCTCCACTCAAC 898

Qy        277 GlyGlyAlaPheCysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuVal 296
Db        899 GGAGGGGCCTTCTGCGAGGGCCAG---GCATTCCAGAAGACCGCCTGCACCACCATCTGC 955

Qy        297 SerValAspGlySerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThr 316
Db        956 CCAGTCGATGGGGCGTGGACGGAGTGGAGCAAGTGGTCAGCCTGCAGCACTGAGTGTGCC 1015

Qy        317 HisTrpArgSerArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGln 336
Db       1016 CACTGGCGTAGCCGCGAGTGCATGGCGCCCCCACCCCAGAACGGAGGCCGTGACTGCAGC 1075

Qy        337 GlyThrAspLeuAspThrArgAsnCysThrSerAspLeuCysValHis---SerAlaSer 355
Db       1076 GGGACGCTGCTCGACTCTAAGAACTGCACAGATGGGCTGTGCATGCAACTGGAGGCCTCA 1135
```

```
Qy        356 GlyProGluAspValAlaLeuTyrValGlyLeu---IleAlaValAlaValCysLeuVal 374
Db       1136 GGG------GATGCGGCGCTGTATGCGGGGCTCGTGGTGGCCATCTTCGTGGTCGTGGCA 1189

Qy        375 LeuLeuLeuLeuValLeuIleLeuValTyrCysArgLysLysGluGlyLeuAspSerAsp 394
Db       1190 ATCCTCATGGCGGTGGGGGTGGTGGTGTACCGCCGCAACTGCCGTGACTTCGACACAGAC 1249

Qy        395 ValAlaAspSerSer---IleLeuThrSerGlyPheGlnProValSerIleLysProSer 413
Db       1250 ATCACTGACTCATCTGCTGCCCTGACTGGTGGTTTCCACCCCGTCAACTTTAAGACGGCA 1309

Qy        414 LysAlaAspAsnProHisLeuLeu------ThrIleGlnProAspLeuSerThrThrThr 431
Db       1310 AGGCCCAGTAACCCGCAGCTCCTACACCCCTCTGTGCCTCCTGACCTGACAGCCAGCGCC 1369

Qy        432 ---ThrTyrGlnGlySerLeuCysProArgGlnAspGlyProSerProLysPheGlnLeu 450
Db       1370 GGCATCTACCGCGGACCCGTGTATGCCCTGCAGGACTCC---ACCGACAAAATCCCCATG 1426

Qy        451 ThrAsnGlyHisLeuLeuSerProLeuGlyGlyGlyArgHisThrLeuHisHisSerSer 470
Db       1427 ACCAACTCTCCTCTGCTGGACCCCTTACCCAGCCTTAAGGTCAAGGTCTACAGCTCCAGC 1486

Qy        471 ProThr-----------------SerGluAlaGluGluPheValSerArgLeuSerThr 484
Db       1487 ACCACGGGCTCTGGGCCAGGCCTGGCAGATGGGGCTGACCTGCTGGGGGTCTTGCCGCCT 1546

Qy        485 GlnAsnTyr-----------------------------------PheArgSer------ 490
Db       1547 GGCACATACCCTAGCGATTTCGCCCGGGACACCCACTTCCTGCACCTGCGCAGCGCCAGC 1606

Qy        491 -----------------------LeuProArgGlyThrSerAsnMetThrTyrGlyThr 502
Db       1607 CTCGGTTCCCAGCAGCTCTTGGGCCTGCCCCGAGACCCAGGGAGCAGCGTCAGCGGCACC 1666

Qy        503 PheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeuLeuIlePro 522
Db       1667 TTTGGCTGCCTGGGTGGGAGGCTCAGCATCCCCGGCACAGGGGTCAGCTTGCTGGTGCCC 1726

Qy        523 ProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHisLysProGlu 542
Db       1727 AATGGAGCCATTCCCCAGGGCAAGTTCTACGAGATGTATCTACTCATCAACAAGGCAGAA 1786

Qy        543 AspValArgLeuProLeuAla---GlyCysGlnThrLeuLeuSerProIleValSerCys 561
Db       1787 AGTACC---CTGCCGCTTTCAGAAGGGACCCAGACAGTATTGAGCCCCTCGGTGACCTGT 1843

Qy        562 GlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGlyGlu 581
Db       1844 GGACCCACAGGCCTCCTGCTGTGCCGCCCCGTCATCCTCACCATGCCCCACTGTGCCGAA 1903

Qy        582 ProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySerTrpGlu 601
Db       1904 GTCAGTGCCCGTGACTGGATCTTTCAGCTCAAGACCCAGGCCCACCAGGGCCACTGGGAG 1963

Qy        602 GlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGlnLeuGlu 621
Db       1964 ---GAGGTGGTGACCCTGGATGAGGAGACCCTGAACACACCCTGCTACTGCCAGCTGGAG 2020

Qy        622 AlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGluAla 641
Db       2021 CCCAGGGCCTGTCACATCCTGCTGGACCAGCTGGGCACCTACGTGTTCACGGGCGAGTCC 2080

Qy        642 LeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAlaCysThr 661
Db       2081 TATTCCCGCTCAGCAGTCAAGCGGCTCCAGCTGGCCGTCTTCGCCCCCGCCCTCTGCACC 2140

Qy        662 SerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuLysGlu 681
Db       2141 TCCCTGGAGTACAGCCTCCGGGTCTACTGCCTGGAGGACACGCCTGTAGCACTGAAGGAG 2200

Qy        682 ValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgValLeuHis 701
Db       2201 GTGCTGGAGCTGGAGCGGACTCTGGGCGGATACTTGGTGGAGGAGCCGAAACCGCTAATG 2260

Qy        702 PheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSerLeuTrp 721
Db       2261 TTCAAGGACAGTTACCACAACCTGCGCCTCTCCCTCCATGACCTCCCCCATGCCCATTGG 2320

Qy        722 LysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsnGlyThr 741
Db       2321 AGGAGCAAGCTGCTGGCCAAATACCAGGAGATCCCCTTCTATCACATTTGGAGTGGCAGC 2380

Qy        742 GlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSerAspLeu 761
Db       2381 CAGAAGGCCCTCCACTGCACTTTCACCCTGGAGAGGCACAGCTTGGCCTCCACAGAGCTC 2440

Qy        762 AlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIleAsnPhe 781
Db       2441 ACCTGCAAGATCTGCGTGCGGCAAGTGGAAGGGGAGGGCCAGATATTCCAGCTGCATACC 2500

Qy        782 AsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAlaGly--- 800
Db       2501 ACTCTGGCA---GAGACACCTGCTGGCTCCCTGGACACTCTCTGCTCTGCCCCTGGCAGC 2557

Qy        801 ---ValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
Db       2558 ACTGTCACCACCCAGCTGGGACCTTATGCCTTCAAGATCCCACTGTCCATCCGCCAGAAG 2617

Qy        820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839
Db       2618 ATATGCAACAGCCTAGATGCCCCCAACTCACGGGGCAATGACTGGCGGATGTTAGCACAG 2677

Qy        840 LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859
Db       2678 AAGCTCTCTATGGACCGGTACCTGAATTACTTTGCCACCAAAGCGAGCCCCACGGGTGTG 2737

Qy        860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879
Db       2738 ATCCTGGACCTCTGGGAAGCTCTGCAGCAGGACGATGGGGACCTCAACAGCCTGGCGAGT 2797

Qy        880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
Db       2798 GCCTTGGAGGAGATGGGCAAGAGTGAGATGCTGGTGGCTGTGGCCACCGACGGGGACTGC 2857
```

```
RESULT 13
AAS21316
ID   AAS21316 standard; cDNA; 3884 BP.
XX
AC   AAS21316;
XX
DT   24-OCT-2001  (first entry)
XX
DE   Human cDNA sequence encoding for PRO4326 polypeptide.
XX
KW   Human secretory and transmembrane; PRO; mammalian; cancer; lung; breast;
KW   prostate; cervical; tumour necrosis factor-alpha; TNF-alpha; cartilage;
KW   ear; proliferation; glucose; free fatty acid; skeletal muscle; adipocyte;
KW   A-peptide; factor VIIA; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   WO200140466-A2.
XX
PD   07-JUN-2001.
XX
PF   01-DEC-2000; 2000WO-US032678.
XX
PR   01-DEC-1999;   99WO-US028301.
PR   01-DEC-1999;   99WO-US028634.
PR   02-DEC-1999;   99WO-US028551.
PR   02-DEC-1999;   99WO-US028564.
PR   02-DEC-1999;   99WO-US028565.
PR   09-DEC-1999;   99US-0170262P.
```

```
PR   16-DEC-1999;   99WO-US030095.
PR   20-DEC-1999;   99WO-US030911.
PR   20-DEC-1999;   99WO-US030999.
PR   30-DEC-1999;   99WO-US031243.
PR   30-DEC-1999;   99WO-US031274.
PR   05-JAN-2000; 2000WO-US000219.
PR   06-JAN-2000; 2000WO-US000277.
PR   06-JAN-2000; 2000WO-US000376.
PR   11-FEB-2000; 2000WO-US003565.
PR   18-FEB-2000; 2000WO-US004341.
PR   18-FEB-2000; 2000WO-US004342.
PR   22-FEB-2000; 2000WO-US004414.
PR   24-FEB-2000; 2000WO-US004914.
PR   24-FEB-2000; 2000WO-US005004.
PR   01-MAR-2000; 2000WO-US005601.
PR   02-MAR-2000; 2000WO-US005841.
PR   03-MAR-2000; 2000US-0187202P.
PR   10-MAR-2000; 2000WO-US006319.
PR   15-MAR-2000; 2000WO-US006884.
PR   20-MAR-2000; 2000WO-US007377.
PR   21-MAR-2000; 2000WO-US007532.
PR   30-MAR-2000; 2000WO-US008439.
PR   17-MAY-2000; 2000WO-US013705.
PR   22-MAY-2000; 2000WO-US014042.
PR   30-MAY-2000; 2000WO-US014941.
PR   02-JUN-2000; 2000WO-US015264.
PR   05-JUN-2000; 2000US-0209832P.
PR   28-JUL-2000; 2000WO-US020710.
PR   11-AUG-2000; 2000WO-US022031.
PR   23-AUG-2000; 2000WO-US023522.
PR   24-AUG-2000; 2000WO-US023328.
PR   08-NOV-2000; 2000WO-US030952.
PR   10-NOV-2000; 2000WO-US030873.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Baker KP,  Beresini M,  Deforge L,  Desnoyers L,  Filvaroff E,  Gao W;
PI   Gerritsen ME,  Goddard A,  Godowski PJ,  Gurney AL,  Sherwood S;
PI   Smith V,  Stewart TA,  Tumas D,  Watanabe CK,  Wood WI,  Zhang Z;
XX
DR   WPI; 2001-408281/43.
DR   P-PSDB; AAU12244.
XX
PT   Isolated , secretory and transmembrane PRO polypeptide used to detect
PT   other PRO polypeptides, link bioactive molecules to cells expressing PRO
PT   polypeptides, and detect the presence of mammalian tumors e.g. lung,
PT   breast, prostate, cervical.
XX
PS   Claim 3; Fig 145; 813pp; English.
XX
CC   AAS21244-AAS21518 encode for novel human secretory and transmembrane PRO
CC   polypeptides. The PRO polypeptides are useful to detect other PRO
CC   polypeptides, to link bioactive molecules to cells expressing PRO
CC   polypeptides, to modulate biological activities of cells expressing PRO
CC   polypeptides, and to detect the presence of mammalian lung, colon,
CC   breast, prostate, rectal, cervical or liver tumours by comparing PRO
CC   polypeptide expression in a cell sample to that in a control sample. Some
CC   of the 275 sequences are also useful to stimulate the release of tumour
CC   necrosis factor-alpha (TNF-alpha) from human blood, the proliferation or
CC   differentiation of chondrocytes, the proliferation or gene expression in
CC   pericyte cells, the release of proteoglycans from cartilage, the
CC   proliferation of inner ear utricular supporting cells or of T-
CC   lymphocytes, the release of a cytokine from peripheral blood monocytes
CC   (PBMCs), or the proliferation of endothelial cells. Some of the PRO
CC   polypeptides may modulate glucose or free fatty acid uptake by skeletal
CC   muscle cells or by adipocytes; or inhibit binding of A-peptide to factor
CC   VIIA. The PRO polypeptides can be used in assays to identify molecules
CC   involved in binding interactions. The polynucleotides encoding PRO
CC   polypeptides can be used to generate probes, antisense RNA/DNA,
CC   transgenic or knock out animals and can be used in gene therapy
XX
SQ   Sequence 3884 BP; 767 A; 1278 C; 1162 G; 677 T; 0 U; 0 Other;
```

```
Alignment Scores:
Pred. No.:              2.96e-161      Length:        3884
Score:                  2479.00        Matches:       493
Percent Similarity:     67.54%         Conservative:  150
Best Local Similarity:  51.79%         Mismatches:    249
Query Match:            51.79%         Indels:        60
DB:                     4              Gaps:          18

US-09-970-944-2 (1-899) x AAS21316 (1-3884)

Qy      1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrp--- 19
Db    398 ATGGGGGCCCGGAGCGGAGCTCGGGGCGCGCTGCTGCTGGCACTGCTGCTCTGCTGGGAC 457

Qy     20 ------LeuArgGlySerGlyAlaGlnGln---SerAlaThrValAlaAsnProValPro 36
Db    458 CCGAGGCTGAGCCAAGCAGGCACTGATTCTGGCAGCGAGGTGCTCCCTGACTCCTTCCCG 517

Qy     37 GlyAlaAsnProAspLeuLeuProHisPheLeuValGluProGluAspValTyrIleVal 56
Db    518 TCAGCGCCAGCAGAGCCGCTGCCCTACTTCCTGCAGGAGCCACAGGACGCCTACATTGTG 577

Qy     57 LysAsnLysProValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLys 76
Db    578 AAGAACAAGCCTGTGGAGCTCCGCTGCCGCGCCTTCCCCGCCACACAGATCTACTTCAAG 637

Qy     77 CysAsnGlyGluTrpValArgGlnValAspHisValIleGluArgSerThrAspGlySer 96
Db    638 TGCAACGGCGAGTGGGTCAGCCAGAACGACCACGTCACACAGGAAGGCCTGGATGAGGCC 697

Qy     97 SerGlyGluProThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysVal 116
Db    698 ACCGGCCTGCGGGTGCGCGAGGTGCAGATCGAGGTGTCGCGGCAGCAGGTGGAGGAGCTC 757

Qy    117 PheGlyLeuGluGluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLys 136
Db    758 TTTGGGCTGGAGGATTACTGGTGCCAGTGCGTGGCCTGGAGCTCCGCAGGCACCACCAAG 817

Qy    137 SerGlnLysAlaTyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeu 156
Db    818 AGTCGCCGAGCCTACGTCCGCATCGCCTACCTGCGCAAGAACTTCGATCAGGAGCCTCTG 877

Qy    157 AlaLysGluValSerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIle 176
Db    878 GGCAAGGAGGTGCCCCTGGACCATGAGGTTCTCCTGCAGTGCCGCCCGCCGGAGGGGGTG 937

Qy    177 ProProAlaGluValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspPro 196
Db    938 CCTGTGGCCGAGGTGGAATGGCTCAAGAATGAGGATGTCATCGACCCCACCCAGGACACC 997

Qy    197 AsnValTyrIleThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThr 216
Db    998 AACTTCCTGCTCACCATCGACCACAACCTCATCATCCGCCAGGCCCGCCTGTCGGACACT 1057

Qy    217 AlaAsnTyrThrCysValAlaLysAsnIleValAlaArgArgArgSerAlaSerAlaAla 236
Db   1058 GCCAACTATACCTGCGTGGCCAAGAACATCGTGGCCAAACGCCGGAGCACCACTGCCACC 1117

Qy    237 ValIleValTyrValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAla 256
Db   1118 GTCATCGTCTACGTGAATGGCGGCTGGTCCAGCTGGGCAGAGTGGTCACCCTGCTCCAAC 1177

Qy    257 SerCysGlyArgGlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsn 276
Db   1178 CGCTGTGGCCGAGGCTGGCAGAAGCGCACCCGGACCTGCACCAACCCCGCTCCACTCAAC 1237

Qy    277 GlyGlyAlaPheCysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuVal 296
Db   1238 GGAGGGGCCTTCTGCGAGGGCCAG---GCATTCCAGAAGACCGCCTGCACCACCATCTGC 1294

Qy    297 SerValAspGlySerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThr 316
Db   1295 CCAGTCGATGGGGCGTGGACGGAGTGGAGCAAGTGGTCAGCCTGCAGCACTGAGTGTGCC 1354
```

```
Qy    317 HisTrpArgSerArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGln 336
Db   1355 CACTGGCGTAGCCGCGAGTGCATGGCGCCCCCACCCCAGAACGGAGGCCGTGACTGCAGC 1414

Qy    337 GlyThrAspLeuAspThrArgAsnCysThrSerAspLeuCysValHisSer--------- 353
Db   1415 GGGACGCTGCTCGACTCTAAGAACTGCACAGATGGGCTGTGCATGCAAAATAAGAAAACT 1474

Qy    354 ------------------------------AlaSerGlyProGluAspValAlaLeuTyr 363
Db   1475 CTAAGCGACCCCAACAGCCACCTGCTGGAGGCCTCAGGG------GATGCGGCGCTGTAT 1528

Qy    364 ValGlyLeu---IleAlaValAlaValCysLeuValLeuLeuLeuLeuValLeuIleLeu 382
Db   1529 GCGGGGCTCGTGGTGGCCATCTTCGTGGTCGTGGCAATCCTCATGGCGGTGGGGGTGGTG 1588

Qy    383 ValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSer---IleLeu 401
Db   1589 GTGTACCGCCGCAACTGCCGTGACTTCGACACAGACATCACTGACTCATCTGCTGCCCTG 1648

Qy    402 ThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeuLeu 421
Db   1649 ACTGGTGGTTTCCACCCCGTCAACTTTAAGACGGCAAGGCCCAGCAACCCGCAGCTCCTA 1708

Qy    422 ------ThrIleGlnProAspLeuSerThrThrThr---ThrTyrGlnGlySerLeuCys 438
Db   1709 CACCCCTCTGTGCCTCCTGACCTGACAGCCAGCGCCGGCATCTACCGCGGACCCGTGTAT 1768

Qy    439 ProArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerPro 458
Db   1769 GCCCTGCAGGACTCC---ACCGACAAAATCCCCATGACCAACTCTCCTCTGCTGGACCCC 1825

Qy    459 LeuGlyGlyGlyArgHisThrLeuHisHisSerSerProThr------------------ 472
Db   1826 TTACCCAGCCTTAAGGTCAAGGTCTACAGCTCCAGCACCACGGGCTCTGGGCCAGGCCTG 1885

Qy    473 SerGluAlaGluGluPheValSerArgLeuSerThrGlnAsnTyr--------------- 487
Db   1886 GCAGATGGGGCTGACCTGCTGGGGGTCTTGCCGCCTGGCACATACCCTAGCGATTTCGCC 1945

Qy    488 ---------------------PheArgSer----------------------------- 490
Db   1946 CGGGACACCCACTTCCTGCACCTGCGCAGCGCCAGCCTCGGTTCCCAGCAGCTCTTGGGC 2005

Qy    491 LeuProArgGlyThrSerAsnMetThrTyrGlyThrPheAsnPheLeuGlyGlyArgLeu 510
Db   2006 CTGCCCCGAGACCCAGGGAGCAGCGTCAGCGGCACCTTTGGCTGCCTGGGTGGGAGGCTC 2065

Qy    511 MetIleProAsnThrGlyIleSerLeuLeuIleProProAspAlaIleProArgGlyLys 530
Db   2066 AGCATCCCCGGCACAGGGGTCAGCTTGCTGGTGCCCAATGGAGCCATTCCCCAGGGCAAG 2125

Qy    531 IleTyrGluIleTyrLeuThrLeuHisLysProGluAspValArgLeuProLeuAla--- 549
Db   2126 TTCTACGAGATGTATCTACTCATCAACAAGGCAGAAAGTACC---CTCCCGCTTTCAGAA 2182

Qy    550 GlyCysGlnThrLeuLeuSerProIleValSerCysGlyProProGlyValLeuLeuThr 569
Db   2183 GGGACCCAGACAGTATTGAGCCCCTCGGTGACCTGTGGACCCACAGGCCTCCTGCTGTGC 2242

Qy    570 ArgProValIleLeuAlaMetAspHisCysGlyGluProSerProAspSerTrpSerLeu 589
Db   2243 CGCCCCGTCATCCTCACCATGCCCCACTGTGCCGAAGTCAGTGCCCGTGACTGGATCTTT 2302

Qy    590 ArgLeuLysLysGlnSerCysGluGlySerTrpGluGlnAspValLeuHisLeuGlyGlu 609
Db   2303 CAGCTCAAGACCCAGGCCCACCAGGGCCACTGGGAG---GAGGTGGTGACCCTGGATGAG 2359

Qy    610 GluAlaProSerHisLeuTyrTyrCysGlnLeuGluAlaSerAlaCysTyrValPheThr 629
Db   2360 GAGACCCTGAACACACCCTGCTACTGCCAGCTGGAGCCCAGGGCCTGTCACATCCTGCTG 2419

Qy    630 GluGlnLeuGlyArgPheAlaLeuValGlyGluAlaLeuSerValAlaAlaAlaLysArg 649
Db   2420 GACCAGCTGGGCACCTACGTGTTCACGGGCGAGTCCTATTCCCGCTCAGCAGTCAAGCGG 2479

Qy    650 LeuLysLeuLeuLeuPheAlaProValAlaCysThrSerLeuGluTyrAsnIleArgVal 669
Db   2480 CTCCAGCTGGCCGTCTTCGCCCCCGCCCTCTGCACCTCCCTGGAGTACAGCCTCCGGGTC 2539

Qy    670 TyrCysLeuHisAspThrHisAspAlaLeuLysGluValValGlnLeuGluLysGlnLeu 689
Db   2540 TACTGCCTGGAGGACACGCCTGTAGCACTGAAGGAGGTGCTGGAGCTGGAGCGGACTCTG 2599

Qy    690 GlyGlyGlnLeuIleGlnGluProArgValLeuHisPheLysAspSerTyrHisAsnLeu 709
Db   2600 GGCGGATACTTGGTGGAGGAGCCGAAACCGCTAATGTTCAAGGACAGTTACCACAACCTG 2659

Qy    710 ArgLeuSerIleHisAspValProSerSerLeuTrpLysSerLysLeuLeuValSerTyr 729
Db   2660 CGCCTCTCCCTCCATGACCTCCCCCATGCCCATTGGAGGAGCAAGCTGCTGGCCAAATAC 2719

Qy    730 GlnGluIleProPheTyrHisIleTrpAsnGlyThrGlnArgTyrLeuHisCysThrPhe 749
Db   2720 CAGGAGATCCCCTTCTATCACATTTGGAGTGGCAGCCAGAAGGCCCTCCACTGCACTTTC 2779

Qy    750 ThrLeuGluArgValSerProSerThrSerAspLeuAlaCysLysLeuTrpValTrpGln 769
Db   2780 ACCCTGGAGAGGCACAGCTTGGCCTCCACAGAGCTCACCTGCAAGATCTGCGTGCGGCAA 2839

Qy    770 ValGluGlyAspGlyGlnSerPheSerIleAsnPheAsnIleThrLysAspThrArgPhe 789
Db   2840 GTGGAAGGGGAGGGCCAGATATTCCAGCTGCATACCACTCTGGCA---GAGACACCTGCT 2896

Qy    790 AlaGluLeuLeuAlaLeuGluSerGluAlaGly------ValProAlaLeuValGlyPro 807
Db   2897 GGCTCCCTGGACACTCTCTGCTCTGCCCCTGGCAGCACTGTCACCACCCAGCTGGGACCT 2956

Qy    808 SerAlaPheLysIleProPheLeuIleArgGlnLysIleIleSerSerLeuAspProPro 827
Db   2957 TATGCCTTCAAGATCCCACTGTCCATCCGCCAGAAGATATGCAACAGCCTAGATGCCCCC 3016

Qy    828 CysArgArgGlyAlaAspTrpArgThrLeuAlaGlnLysLeuHisLeuAspSerHisLeu 847
Db   3017 AACTCACGGGGCAATGACTGGCGGATGTTAGCACAGAAGCTCTCTATGGACCGGTACCTG 3076

Qy    848 SerPhePheAlaSerLysProSerProThrAlaMetIleLeuAsnLeuTrpGluAlaArg 867
Db   3077 AATTACTTTGCCACCAAAGCGAGCCCCACGGGTGTGATCCTGGACCTCTGGGAAGCTCTG 3136

Qy    868 HisPheProAsnGlyAsnLeuSerGlnLeuAlaAlaAlaValAlaGlyLeuGlyGlnPro 887
Db   3137 CAGCAGGACGATGGGGACCTCAACAGCCTGGCGAGTGCCTTGGAGGAGATGGGCAAGAGT 3196

Qy    888 AspAlaGlyLeuPheThrValSerGluAlaGluCys 899
Db   3197 GAGATGCTGGTGGCTGTGGCCACCGACGGGGACTGC 3232
```

RESULT 14
ACD23925
ID   ACD23925 standard; cDNA; 3884 BP.
XX
AC   ACD23925;
XX
DT   26-AUG-2003  (first entry)
XX
DE   Novel human secreted and transmembrane protein PRO4326 cDNA.
XX
KW   Human; secreted and transmembrane protein; PRO; antiinflammatory;
KW   antiarteriosclerotic; cardiant; anti-infertility; anti-HIV; cytostatic;
KW   antidiabetic; gene therapy; tumour necrosis factor (TNF)-alpha release;
KW   TNF-alpha release; cell proliferation; cell differentiation;
KW   gene expression modulator; proteoglycan release; cytokine release;
KW   tumour; inflammatory disease; organ failure; atherosclerosis;
KW   cardiac injury; infertility; birth defect; premature aging; AIDS;

```
KW   acquired immunodeficiency syndrome; cancer; diabetic complication;
KW   chromosome mapping; gene mapping; pharmaceutical; diagnostic; biosensor;
KW   bioreactor; tissue typing; gene; ss.
XX
OS   Homo sapiens.
XX
PN   US2003032156-A1.
XX
PD   13-FEB-2003.
XX
PF   06-MAY-2002; 2002US-00140474.
XX
PR   31-MAR-1997;   97WO-US005230.
PR   12-JUN-1998;   98WO-US012456.
PR   14-JUL-1998;   98WO-US014552.
PR   28-AUG-1998;   98WO-US017888.
PR   10-SEP-1998;   98WO-US018824.
PR   14-SEP-1998;   98WO-US019093.
PR   14-SEP-1998;   98WO-US019094.
PR   14-SEP-1998;   98WO-US019177.
PR   16-SEP-1998;   98WO-US019330.
PR   17-SEP-1998;   98WO-US019437.
PR   07-OCT-1998;   98WO-US021141.
PR   29-OCT-1998;   98WO-US022991.
PR   29-OCT-1998;   98WO-US022992.
PR   20-NOV-1998;   98WO-US024855.
PR   01-DEC-1998;   98WO-US025108.
PR   05-JAN-1999;   99WO-US000106.
PR   08-MAR-1999;   99WO-US005028.
PR   10-MAR-1999;   99WO-US005190.
PR   20-APR-1999;   99WO-US008615.
PR   14-MAY-1999;   99WO-US010733.
PR   02-JUN-1999;   99WO-US012252.
PR   01-SEP-1999;   99WO-US020111.
PR   08-SEP-1999;   99WO-US020594.
PR   13-SEP-1999;   99WO-US020944.
PR   15-SEP-1999;   99WO-US021090.
PR   15-SEP-1999;   99WO-US021547.
PR   05-OCT-1999;   99WO-US023089.
PR   29-NOV-1999;   99WO-US028214.
PR   30-NOV-1999;   99WO-US028313.
PR   30-NOV-1999;   99WO-US028409.
PR   01-DEC-1999;   99WO-US028301.
PR   01-DEC-1999;   99WO-US028634.
PR   02-DEC-1999;   99WO-US028551.
PR   02-DEC-1999;   99WO-US028564.
PR   02-DEC-1999;   99WO-US028565.
PR   16-DEC-1999;   99WO-US030095.
PR   20-DEC-1999;   99WO-US030911.
PR   20-DEC-1999;   99WO-US030999.
PR   22-DEC-1999;   99WO-US030720.
PR   30-DEC-1999;   99WO-US031243.
PR   30-DEC-1999;   99WO-US031274.
PR   05-JAN-2000; 2000WO-US000219.
PR   06-JAN-2000; 2000WO-US000277.
PR   06-JAN-2000; 2000WO-US000376.
PR   11-FEB-2000; 2000WO-US003565.
PR   18-FEB-2000; 2000WO-US004341.
PR   18-FEB-2000; 2000WO-US004342.
PR   22-FEB-2000; 2000WO-US004414.
PR   24-FEB-2000; 2000WO-US004914.
PR   24-FEB-2000; 2000WO-US005004.
PR   01-MAR-2000; 2000WO-US005601.
PR   02-MAR-2000; 2000WO-US005746.
PR   02-MAR-2000; 2000WO-US005841.
PR   10-MAR-2000; 2000WO-US006319.
PR   15-MAR-2000; 2000WO-US006884.
PR   20-MAR-2000; 2000WO-US007377.
PR   21-MAR-2000; 2000WO-US007532.
PR   30-MAR-2000; 2000WO-US008439.
PR   17-MAY-2000; 2000WO-US013705.
PR   22-MAY-2000; 2000WO-US014042.
PR   30-MAY-2000; 2000WO-US014941.
PR   02-JUN-2000; 2000WO-US015264.
PR   28-JUL-2000; 2000WO-US020710.
PR   11-AUG-2000; 2000WO-US022031.
PR   23-AUG-2000; 2000WO-US023522.
PR   24-AUG-2000; 2000WO-US023328.
PR   08-NOV-2000; 2000WO-US030952.
PR   10-NOV-2000; 2000WO-US030873.
PR   01-DEC-2000; 2000WO-US032678.
PR   20-DEC-2000; 2000US-00747259.
PR   20-DEC-2000; 2000WO-US034956.
PR   28-FEB-2001; 2001US-00796498.
PR   28-FEB-2001; 2001WO-US006520.
PR   01-MAR-2001; 2001WO-US006666.
PR   09-MAR-2001; 2001US-00802706.
PR   14-MAR-2001; 2001US-00808689.
PR   22-MAR-2001; 2001US-00816744.
PR   05-APR-2001; 2001US-00828366.
PR   10-MAY-2001; 2001US-00854208.
PR   10-MAY-2001; 2001US-00854280.
PR   18-MAY-2001; 2001US-00860216.
PR   25-MAY-2001; 2001US-00866028.
PR   25-MAY-2001; 2001US-00866034.
PR   25-MAY-2001; 2001WO-US017092.
PR   01-JUN-2001; 2001US-00872035.
PR   01-JUN-2001; 2001WO-US017800.
PR   05-JUN-2001; 2001US-00874503.
PR   14-JUN-2001; 2001US-00882636.
PR   19-JUN-2001; 2001US-00886342.
PR   20-JUN-2001; 2001WO-US019692.
PR   21-JUN-2001; 2001US-00887879.
PR   22-JUN-2001; 2001WO-US020116.
PR   29-JUN-2001; 2001WO-US021066.
PR   09-JUL-2001; 2001WO-US021735.
PR   18-JUL-2001; 2001US-00908827.
PR   06-AUG-2001; 2001US-00924419.
PR   09-AUG-2001; 2001US-00927796.
PR   16-AUG-2001; 2001US-00931836.
PR   19-DEC-2001; 2001US-00028072.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Baker KP,  Beresini M,  Deforge L,  Desnoyers L,  Filvaroff E,  Gao W;
PI   Gerritsen ME,  Goddard A,  Godowski PJ,  Gurney AL,  Sherwood S;
PI   Smith V,  Stewart TA,  Tumas D,  Watanabe CK,  Wood WI,  Zhang Z;
XX
DR   WPI; 2003-341980/32.
DR   P-PSDB; ABO17688.
XX
PT   New secreted and transmembrane PRO nucleic acids, for treating
PT   inflammation, organ failure, atherosclerosis, cardiac injury,
PT   infertility, birth defects, premature aging, axquired immunodeficiency
PT   syndrome (AIDS), or cancer.
XX
PS   Claim 2; Fig 145; 660pp; English.
XX
CC   The invention describes an isolated nucleic acid (I) comprising, or which
CC   has 80 % sequence identity to, or the full-length coding sequence of, one
CC   of 275 nucleotide sequences, and which encodes a corresponding
CC   polypeptide selected from 275 amino acid sequences, where all sequences
CC   are given in the specification. The polypeptide encoded by (I) is used to
CC   detect PRO polypeptides, link a bioactive molecule to a cell expressing a
CC   PRO polypeptide, modulate a biological activity of a cell, stimulate the
CC   release of tumour necrosis factor (TNF)-alpha from human blood, modulate
CC   the uptake of glucose or free fatty acid by cells, stimulate or inhibit
CC   the proliferation or differentiation of cells or gene expression,
CC   stimulate the release of proteoglycans, stimulate the release of cytokine
CC   from peripheral blood mononuclear cells, inhibit the binding of A-peptide
CC   to factor VIIA, or detect the presence of tumour in a mammal. The nucleic
CC   acid and polypeptide encoded by it, are useful for treating inflammatory
CC   diseases, organ failure, atherosclerosis, cardiac injury, infertility,
CC   birth defects, premature aging, acquired immunodeficiency syndrome
CC   (AIDS), cancer, or diabetic complications. The nucleic acid is useful as
CC   hybridisation probes, in chromosome and gene mapping, and in generating
```

```
CC   antisense RNA or DNA. The polypeptides are useful as pharmaceuticals,
CC   diagnostics, biosensors or bioreactors. Both are useful in tissue typing.
CC   This sequence encodes a novel human secreted and transmembrane PRO
CC   polypeptide
XX
SQ   Sequence 3884 BP; 767 A; 1278 C; 1162 G; 677 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              2.96e-161     Length:         3884
Score:                  2479.00       Matches:        493
Percent Similarity:     67.54%        Conservative:   150
Best Local Similarity:  51.79%        Mismatches:     249
Query Match:            51.79%        Indels:         60
DB:                     7             Gaps:           18

US-09-970-944-2 (1-899) x ACD23925 (1-3884)

Qy      1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrp--- 19
Db    398 ATGGGGGCCCGGAGCGGAGCTCGGGGCGCGCTGCTGCTGGCACTGCTGCTCTGCTGGGAC 457

Qy     20 ------LeuArgGlySerGlyAlaGlnGln---SerAlaThrValAlaAsnProValPro 36
Db    458 CCGAGGCTGAGCCAAGCAGGCACTGATTCTGGCAGCGAGGTGCTCCCTGACTCCTTCCCG 517

Qy     37 GlyAlaAsnProAspLeuLeuProHisPheLeuValGluProGluAspValTyrIleVal 56
Db    518 TCAGCGCCAGCAGAGCCGCTGCCCTACTTCCTGCAGGAGCCACAGGACGCCTACATTGTG 577

Qy     57 LysAsnLysProValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLys 76
Db    578 AAGAACAAGCCTGTGGAGCTCCGCTGCCGCGCCTTCCCCGCCACACAGATCTACTTCAAG 637

Qy     77 CysAsnGlyGluTrpValArgGlnValAspHisValIleGluArgSerThrAspGlySer 96
Db    638 TGCAACGGCGAGTGGGTCAGCCAGAACGACCACGTCACACAGGAAGGCCTGGATGAGGCC 697

Qy     97 SerGlyGluProThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysVal 116
Db    698 ACCGGCCTGCGGGTGCGCGAGGTGCAGATCGAGGTGTCGCGGCAGCAGGTGGAGGAGCTC 757

Qy    117 PheGlyLeuGluGluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLys 136
Db    758 TTTGGGCTGGAGGATTACTGGTGCCAGTGCGTGGCCTGGAGCTCCGCAGGCACCACCAAG 817

Qy    137 SerGlnLysAlaTyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeu 156
Db    818 AGTCGCCGAGCCTACGTCCGCATCGCCTACCTGCGCAAGAACTTCGATCAGGAGCCTCTG 877

Qy    157 AlaLysGluValSerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIle 176
Db    878 GGCAAGGAGGTGCCCCTGGACCATGAGGTTCTCCTGCAGTGCCGCCCGCCGGAGGGGGTG 937

Qy    177 ProProAlaGluValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspPro 196
Db    938 CCTGTGGCCGAGGTGGAATGGCTCAAGAATGAGGATGTCATCGACCCCACCCAGGACACC 997

Qy    197 AsnValTyrIleThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThr 216
Db    998 AACTTCCTGCTCACCATCGACCACAACCTCATCATCCGCCAGGCCCGCCTGTCGGACACT 1057

Qy    217 AlaAsnTyrThrCysValAlaLysAsnIleValAlaArgArgArgSerAlaSerAlaAla 236
Db   1058 GCCAACTATACCTGCGTGGCCAAGAACATCGTGGCCAAACGCCGGAGCACCACTGCCACC 1117

Qy    237 ValIleValTyrValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAla 256
Db   1118 GTCATCGTCTACGTGAATGGCGGCTGGTCCAGCTGGGCAGAGTGGTCACCCTGCTCCAAC 1177

Qy    257 SerCysGlyArgGlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsn 276
Db   1178 CGCTGTGGCCGAGGCTGGCAGAAGCGCACCCGGACCTGCACCAACCCCGCTCCACTCAAC 1237

Qy    277 GlyGlyAlaPheCysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuVal 296
Db   1238 GGAGGGGCCTTCTGCGAGGGCCAG---GCATTCCAGAAGACCGCCTGCACCACCATCTGC 1294

Qy    297 SerValAspGlySerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThr 316
Db   1295 CCAGTCGATGGGGCGTGGACGGAGTGGAGCAAGTGGTCAGCCTGCAGCACTGAGTGTGCC 1354

Qy    317 HisTrpArgSerArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGln 336
Db   1355 CACTGGCGTAGCCGCGAGTGCATGGCGCCCCCACCCCAGAACGGAGGCCGTGACTGCAGC 1414

Qy    337 GlyThrAspLeuAspThrArgAsnCysThrSerAspLeuCysValHisSer--------- 353
Db   1415 GGGACGCTGCTCGACTCTAAGAACTGCACAGATGGGCTGTGCATGCAAAATAAGAAAACT 1474

Qy    354 -----------------------------AlaSerGlyProGluAspValAlaLeuTyr 363
Db   1475 CTAAGCGACCCCAACAGCCACCTGCTGGAGGCCTCAGGG------GATGCGGCGCTGTAT 1528

Qy    364 ValGlyLeu---IleAlaValAlaValCysLeuValLeuLeuLeuLeuValLeuIleLeu 382
Db   1529 GCGGGGCTCGTGGTGGCCATCTTCGTGGTCGTGGCAATCCTCATGGCGGTGGGGGTGGTG 1588

Qy    383 ValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSer---IleLeu 401
Db   1589 GTGTACCGCCGCAACTGCCGTGACTTCGACACAGACATCACTGACTCATCTGCTGCCCTG 1648

Qy    402 ThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeuLeu 421
Db   1649 ACTGGTGGTTTCCACCCCGTCAACTTTAAGACGGCAAGGCCCAGCAACCCGCAGCTCCTA 1708

Qy    422 ------ThrIleGlnProAspLeuSerThrThrThr---ThrTyrGlnGlySerLeuCys 438
Db   1709 CACCCCTCTGTGCCTCCTGACCTGACAGCCAGCGCCGGCATCTACCGCGGACCCGTGTAT 1768

Qy    439 ProArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerPro 458
Db   1769 GCCCTGCAGGACTCC---ACCGACAAAATCCCCATGACCAACTCTCCTCTGCTGGACCCC 1825

Qy    459 LeuGlyGlyGlyArgHisThrLeuHisHisSerSerProThr------------------ 472
Db   1826 TTACCCAGCCTTAAGGTCAAGGTCTACAGCTCCAGCACCACGGGCTCTGGGCCAGGCCTG 1885

Qy    473 SerGluAlaGluGluPheValSerArgLeuSerThrGlnAsnTyr--------------- 487
Db   1886 GCAGATGGGGCTGACCTGCTGGGGGTCTTGCCGCCTGGCACATACCCTAGCGATTTCGCC 1945

Qy    488 --------------------PheArgSer------------------------------ 490
Db   1946 CGGGACACCCACTTCCTGCACCTGCGCAGCGCCAGCCTCGGTTCCCAGCAGCTCTTGGGC 2005

Qy    491 LeuProArgGlyThrSerAsnMetThrTyrGlyThrPheAsnPheLeuGlyGlyArgLeu 510
Db   2006 CTGCCCCGAGACCCAGGGAGCAGCGTCAGCGGCACCTTTGGCTGCCTGGGTGGGAGGCTC 2065

Qy    511 MetIleProAsnThrGlyIleSerLeuLeuIleProProAspAlaIleProArgGlyLys 530
Db   2066 AGCATCCCCGGCACAGGGGTCAGCTTGCTGGTGCCCAATGGAGCCATTCCCCAGGGCAAG 2125

Qy    531 IleTyrGluIleTyrLeuThrLeuHisLysProGluAspValArgLeuProLeuAla--- 549
Db   2126 TTCTACGAGATGTATCTACTCATCAACAAGGCAGAAAGTACC---CTCCCGCTTTCAGAA 2182

Qy    550 GlyCysGlnThrLeuLeuSerProIleValSerCysGlyProProGlyValLeuLeuThr 569
Db   2183 GGGACCCAGACAGTATTGAGCCCCTCGGTGACCTGTGGACCCACAGGCCTCCTGCTGTGC 2242

Qy    570 ArgProValIleLeuAlaMetAspHisCysGlyGluProSerProAspSerTrpSerLeu 589
Db   2243 CGCCCCGTCATCCTCACCATGCCCCACTGTGCCGAAGTCAGTGCCCGTGACTGGATCTTT 2302

Qy    590 ArgLeuLysLysGlnSerCysGluGlySerTrpGluGlnAspValLeuHisLeuGlyGlu 609
```

```
             :::||||||  |||:::   :::|||  ||||||  :::|||:::  |||  |||
Db     2303 CAGCTCAAGACCCAGGCCCACCAGGGCCACTGGGAG---GAGGTGGTGACCCTGGATGAG 2359

Qy      610 GluAlaProSerHisLeuTyrTyrCysGlnLeuGluAlaSerAlaCysTyrValPheThr 629
             |||      :::          |||||||||||||       ||||||::::::
Db     2360 GAGACCCTGAACACACCCTGCTACTGCCAGCTGGAGCCCAGGGCCTGTCACATCCTGCTG 2419

Qy      630 GluGlnLeuGlyArgPheAlaLeuValGlyGluAlaLeuSerValAlaAlaAlaLysArg 649
             :::|||||||||    :::          ||||||:::   |||   :::|||   ||||||
Db     2420 GACCAGCTGGGCACCTACGTGTTCACGGGCGAGTCCTATTCCCGCTCAGCAGTCAAGCGG 2479

Qy      650 LeuLysLeuLeuLeuPheAlaProValAlaCysThrSerLeuGluTyrAsnIleArgVal 669
             |||:::|||   :::|||||||||     ||||||||||||||||||:::::::||||||
Db     2480 CTCCAGCTGGCCGTCTTCGCCCCCGCCCTCTGCACCTCCCTGGAGTACAGCCTCCGGGTC 2539

Qy      670 TyrCysLeuHisAspThrHisAspAlaLeuLysGluValValGlnLeuGluLysGlnLeu 689
             |||||||||   ||||||     ||||||||||||||||:::::||||||:::   |||
Db     2540 TACTGCCTGGAGGACACGCCTGTAGCACTGAAGGAGGTGCTGGAGCTGGAGCGGACTCTG 2599

Qy      690 GlyGlyGlnLeuIleGlnGluProArgValLeuHisPheLysAspSerTyrHisAsnLeu 709
             ||||||   |||::::::||||||:::   |||   ||||||||||||||||||||||||
Db     2600 GGCGGATACTTGGTGGAGGAGCCGAAACCGCTAATGTTCAAGGACAGTTACCACAACCTG 2659

Qy      710 ArgLeuSerIleHisAspValProSerSerLeuTrpLysSerLysLeuLeuValSerTyr 729
             |||||||||:::||||||:::|||    :::    |||:::||||||||||||     |||
Db     2660 CGCCTCTCCCTCCATGACCTCCCCCATGCCCATTGGAGGAGCAAGCTGCTGGCCAAATAC 2719

Qy      730 GlnGluIleProPheTyrHisIleTrpAsnGlyThrGlnArgTyrLeuHisCysThrPhe 749
             ||||||||||||||||||||||||||||:::|||:::|||:::   |||||||||||||||
Db     2720 CAGGAGATCCCCTTCTATCACATTTGGAGTGGCAGCCAGAAGGCCCTCCACTGCACTTTC 2779

Qy      750 ThrLeuGluArgValSerProSerThrSerAspLeuAlaCysLysLeuTrpValTrpGln 769
             ||||||||||||   |||   ::::::::::::|||   ||||||:::   |||   |||
Db     2780 ACCCTGGAGAGGCACAGCTTGGCCTCCACAGAGCTCACCTGCAAGATCTGCGTGCGGCAA 2839

Qy      770 ValGluGlyAspGlyGlnSerPheSerIleAsnPheAsnIleThrLysAspThrArgPhe 789
             |||||||||:::||||||   |||   ::::::      :::      :::|||
Db     2840 GTGGAAGGGGAGGGCCAGATATTCCAGCTGCATACCACTCTGGCA---GAGACACCTGCT 2896

Qy      790 AlaGluLeuLeuAlaLeuGluSerGluAlaGly------ValProAlaLeuValGlyPro 807
                   |||     |||   |||      |||      |||         :::||||||
Db     2897 GGCTCCCTGGACACTCTCTGCTCTGCCCCTGGCAGCACTGTCACCACCCAGCTGGGACCT 2956

Qy      808 SerAlaPheLysIleProPheLeuIleArgGlnLysIleIleSerSerLeuAspProPro 827
                ||||||||||||||||      |||||||||||||||   :::|||||||||   |||
Db     2957 TATGCCTTCAAGATCCCACTGTCCATCCGCCAGAAGATATGCAACAGCCTAGATGCCCCC 3016

Qy      828 CysArgArgGlyAlaAspTrpArgThrLeuAlaGlnLysLeuHisLeuAspSerHisLeu 847
                   ||||||   |||||||||   |||||||||||||||      :::|||   :::|||
Db     3017 AACTCACGGGGCAATGACTGGCGGATGTTAGCACAGAAGCTCTCTATGGACCGGTACCTG 3076

Qy      848 SerPhePheAlaSerLysProSerProThrAlaMetIleLeuAsnLeuTrpGluAlaArg 867
             ::::::||||||:::|||   |||||||||   :::||||||:::||||||||||||
Db     3077 AATTACTTTGCCACCAAAGCGAGCCCCACGGGTGTGATCCTGGACCTCTGGGAAGCTCTG 3136

Qy      868 HisPheProAsnGlyAsnLeuSerGlnLeuAlaAlaAlaValAlaGlyLeuGlyGlnPro 887
                      :::|||:::|||:::   ||||||:::|||:::      :::|||:::
Db     3137 CAGCAGGACGATGGGGACCTCAACAGCCTGGCGAGTGCCTTGGAGGAGATGGGCAAGAGT 3196

Qy      888 AspAlaGlyLeuPheThrValSerGluAlaGluCys 899
             :::      :::      ::::::   :::|||
Db     3197 GAGATGCTGGTGGCTGTGGCCACCGACGGGGACTGC 3232
```

```
RESULT 15
ACA67066
ID   ACA67066 standard; cDNA; 3884 BP.
XX
AC   ACA67066;
XX
DT   23-JUN-2003  (first entry)
XX
DE   cDNA encoding human PRO polypeptide #73.
XX
KW   Human; PRO polypeptide; secreted and transmembrane protein;
KW   anti-PRO antibody; diagnostic assay; gene expression; diabetes;
KW   bone disorder; cartilage disorder; rheumatoid arthritis; obesity;
KW   sports injury; osteoarthritis; hyper-insulinaemia; hypo-insulinaemia;
KW   hearing loss; coagulation disorder; stroke; heart attack; cardiant;
KW   antidiabetic; anorectic; vulnerary; antiarthritic; osteopathic;
KW   antirheumatic; auditory; cerebroprotective; angiogenic; gene; ss.
XX
OS   Homo sapiens.
XX
PN   US2003004311-A1.
XX
PD   02-JAN-2003.
XX
PF   19-DEC-2001; 2001US-00028072.
XX
PR   18-JUN-1997;   97US-0049911P.
PR   26-AUG-1997;   97US-0056974P.
PR   17-SEP-1997;   97US-0059113P.
PR   17-SEP-1997;   97US-0059115P.
PR   17-SEP-1997;   97US-0059117P.
PR   17-SEP-1997;   97US-0059122P.
PR   17-SEP-1997;   97US-0059184P.
PR   18-SEP-1997;   97US-0059263P.
PR   19-SEP-1997;   97US-0059352P.
PR   19-SEP-1997;   97US-0059588P.
PR   24-SEP-1997;   97US-0059836P.
PR   17-OCT-1997;   97US-0062250P.
PR   17-OCT-1997;   97US-0062285P.
PR   17-OCT-1997;   97US-0062287P.
PR   17-OCT-1997;   97US-0063755P.
PR   24-OCT-1997;   97US-0062814P.
PR   24-OCT-1997;   97US-0062816P.
PR   24-OCT-1997;   97US-0063045P.
PR   24-OCT-1997;   97US-0063082P.
PR   24-OCT-1997;   97US-0063127P.
PR   27-OCT-1997;   97US-0063327P.
PR   27-OCT-1997;   97US-0063329P.
PR   28-OCT-1997;   97US-0063550P.
PR   28-OCT-1997;   97US-0063561P.
PR   29-OCT-1997;   97US-0063704P.
PR   29-OCT-1997;   97US-0063733P.
PR   29-OCT-1997;   97US-0063735P.
PR   29-OCT-1997;   97US-0063738P.
PR   03-NOV-1997;   97US-0064248P.
PR   07-NOV-1997;   97US-0064809P.
PR   12-NOV-1997;   97US-0065186P.
PR   17-NOV-1997;   97US-0065846P.
PR   21-NOV-1997;   97US-0066364P.
PR   24-NOV-1997;   97US-0066453P.
PR   24-NOV-1997;   97US-0066511P.
PR   24-NOV-1997;   97US-0066770P.
PR   11-DEC-1997;   97US-0069212P.
PR   11-DEC-1997;   97US-0069278P.
PR   11-DEC-1997;   97US-0069334P.
PR   16-DEC-1997;   97US-0069694P.
PR   23-JAN-1998;   98US-0072320P.
PR   04-FEB-1998;   98US-0073612P.
PR   09-FEB-1998;   98US-0074086P.
PR   09-FEB-1998;   98US-0074092P.
PR   12-MAR-1998;   98US-0077791P.
PR   20-MAR-1998;   98US-0078910P.
PR   25-MAR-1998;   98US-0079294P.
PR   27-MAR-1998;   98US-0079663P.
PR   27-MAR-1998;   98US-0079728P.
PR   31-MAR-1998;   98US-0080165P.
PR   12-JUN-1998;   98WO-US012456.
PR   14-JUL-1998;   98WO-US014552.
PR   28-AUG-1998;   98WO-US017888.
PR   10-SEP-1998;   98WO-US018824.
PR   14-SEP-1998;   98WO-US019093.
PR   14-SEP-1998;   98WO-US019094.
PR   14-SEP-1998;   98WO-US019177.
```

```
PR   16-SEP-1998;    98WO-US019330.
PR   17-SEP-1998;    98WO-US019437.
PR   07-OCT-1998;    98WO-US021141.
PR   29-OCT-1998;    98WO-US022991.
PR   29-OCT-1998;    98WO-US022992.
PR   20-NOV-1998;    98WO-US024855.
PR   01-DEC-1998;    98WO-US025108.
PR   05-JAN-1999;    99WO-US000106.
PR   08-MAR-1999;    99WO-US005028.
PR   10-MAR-1999;    99WO-US005190.
PR   20-APR-1999;    99WO-US008615.
PR   14-MAY-1999;    99WO-US010733.
PR   02-JUN-1999;    99WO-US012252.
PR   01-SEP-1999;    99WO-US020111.
PR   08-SEP-1999;    99WO-US020594.
PR   13-SEP-1999;    99WO-US020944.
PR   15-SEP-1999;    99WO-US021090.
PR   15-SEP-1999;    99WO-US021547.
PR   05-OCT-1999;    99WO-US023089.
PR   29-NOV-1999;    99WO-US028214.
PR   30-NOV-1999;    99WO-US028313.
PR   30-NOV-1999;    99WO-US028409.
PR   01-DEC-1999;    99WO-US028301.
PR   01-DEC-1999;    99WO-US028634.
PR   02-DEC-1999;    99WO-US028551.
PR   02-DEC-1999;    99WO-US028564.
PR   02-DEC-1999;    99WO-US028565.
PR   16-DEC-1999;    99WO-US030095.
PR   20-DEC-1999;    99WO-US030911.
PR   20-DEC-1999;    99WO-US030999.
PR   30-DEC-1999;    99WO-US031243.
PR   30-DEC-1999;    99WO-US031274.
PR   05-JAN-2000; 2000WO-US000219.
PR   06-JAN-2000; 2000WO-US000277.
PR   06-JAN-2000; 2000WO-US000376.
PR   11-FEB-2000; 2000WO-US003565.
PR   18-FEB-2000; 2000WO-US004341.
PR   18-FEB-2000; 2000WO-US004342.
PR   22-FEB-2000; 2000WO-US004414.
PR   24-FEB-2000; 2000WO-US004914.
PR   24-FEB-2000; 2000WO-US005004.
PR   01-MAR-2000; 2000WO-US005601.
PR   02-MAR-2000; 2000WO-US005746.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Baker KP,  Beresini M,  Deforge L,  Desnoyers L,  Filvaroff E,  Gao W;
PI   Gerritsen ME,  Goddard A,  Godowski PJ,  Gurney AL,  Sherwood S;
PI   Smith V,  Stewart TA,  Tumas D,  Watanabe CK,  Wood WI,  Zhang Z;
XX
DR   WPI; 2003-352836/33.
DR   P-PSDB; ABU80942.
XX
PT   New isolated PRO polypeptide useful for treating diabetes, rheumatoid
PT   arthritis, sports injuries, obesity, hearing loss in mammals, stroke, or
PT   heart attack.
XX
PS   Claim 2; Fig 145; 643pp; English.
XX
CC   The present invention relates to the isolation of novel human PRO
CC   polypeptides, and the polynucleotide sequences encoding them. The PRO
CC   polypeptides are secreted and transmembrane proteins. The PRO
CC   polypeptides and polynucleotides are useful for preparing a medicament
CC   useful in the treatment of diabetes, bone and/or cartilage disorders
CC   (e.g. rheumatoid arthritis, sports injuries, osteoarthritis), obesity,
CC   hyper- or hypo-insulinaemia, hearing loss, and coagulation disorders
CC   (e.g. stroke, heart attack). Anti-PRO antibodies are useful in diagnostic
CC   assays for PRO, by detecting its expression in specific cells, tissues or
CC   serum, and for affinity purification of PRO from recombinant cell culture
CC   or natural sources. ACA66994-ACA67268 represent cDNA sequences encoding
CC   the human PRO polypeptides of the invention. Note: The sequence data for
CC   this patent was obtained in electronic format directly from the USPTO web
CC   site at seqdata.uspto.gov/psipsDIDEntry.html
XX
SQ   Sequence 3884 BP; 767 A; 1278 C; 1162 G; 677 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              2.96e-161     Length:        3884
Score:                  2479.00       Matches:       493
Percent Similarity:     67.54%        Conservative:  150
Best Local Similarity:  51.79%        Mismatches:    249
Query Match:            51.79%        Indels:        60
DB:                     7             Gaps:          18

US-09-970-944-2 (1-899) x ACA67066 (1-3884)

Qy       1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrp--- 19
Db     398 ATGGGGGCCCGGAGCGGAGCTCGGGGCGCGCTGCTGCTGGCACTGCTGCTCTGCTGGGAC 457

Qy      20 ------LeuArgGlySerGlyAlaGlnGln---SerAlaThrValAlaAsnProValPro 36
Db     458 CCGAGGCTGAGCCAAGCAGGCACTGATTCTGGCAGCGAGGTGCTCCCTGACTCCTTCCCG 517

Qy      37 GlyAlaAsnProAspLeuLeuProHisPheLeuValGluProGluAspValTyrIleVal 56
Db     518 TCAGCGCCAGCAGAGCCGCTGCCCTACTTCCTGCAGGAGCCACAGGACGCCTACATTGTG 577

Qy      57 LysAsnLysProValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLys 76
Db     578 AAGAACAAGCCTGTGGAGCTCCGCTGCCGCGCCTTCCCCGCCACACAGATCTACTTCAAG 637

Qy      77 CysAsnGlyGluTrpValArgGlnValAspHisValIleGluArgSerThrAspGlySer 96
Db     638 TGCAACGGCGAGTGGGTCAGCCAGAACGACCACGTCACACAGGAAGGCCTGGATGAGGCC 697

Qy      97 SerGlyGluProThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysVal 116
Db     698 ACCGGCCTGCGGGTGCGCGAGGTGCAGATCGAGGTGTCGCGGCAGCAGGTGGAGGAGCTC 757

Qy     117 PheGlyLeuGluGluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLys 136
Db     758 TTTGGGCTGGAGGATTACTGGTGCCAGTGCGTGGCCTGGAGCTCCGCAGGCACCACCAAG 817

Qy     137 SerGlnLysAlaTyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeu 156
Db     818 AGTCGCCGAGCCTACGTCCGCATCGCCTACCTGCGCAAGAACTTCGATCAGGAGCCTCTG 877

Qy     157 AlaLysGluValSerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIle 176
Db     878 GGCAAGGAGGTGCCCCTGGACCATGAGGTTCTCCTGCAGTGCCGCCCGCCGGAGGGGGTG 937

Qy     177 ProProAlaGluValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspPro 196
Db     938 CCTGTGGCCGAGGTGGAATGGCTCAAGAATGAGGATGTCATCGACCCCACCCAGGACACC 997

Qy     197 AsnValTyrIleThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThr 216
Db     998 AACTTCCTGCTCACCATCGACCACAACCTCATCATCCGCCAGGCCCGCCTGTCGGACACT 1057

Qy     217 AlaAsnTyrThrCysValAlaLysAsnIleValAlaArgArgArgSerAlaSerAlaAla 236
Db    1058 GCCAACTATACCTGCGTGGCCAAGAACATCGTGGCCAAACGCCGGAGCACCACTGCCACC 1117

Qy     237 ValIleValTyrValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAla 256
Db    1118 GTCATCGTCTACGTGAATGGCGGCTGGTCCAGCTGGGCAGAGTGGTCACCCTGCTCCAAC 1177

Qy     257 SerCysGlyArgGlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsn 276
Db    1178 CGCTGTGGCCGAGGCTGGCAGAAGCGCACCCGGACCTGCACCAACCCCGCTCCACTCAAC 1237

Qy     277 GlyGlyAlaPheCysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuVal 296
Db    1238 GGAGGGGCCTTCTGCGAGGGCCAG---GCATTCCAGAAGACCGCCTGCACCACCATCTGC 1294
```

```
Qy        297 SerValAspGlySerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThr 316
                 ||||||||:::|||:::    |||||||||||||||||||         :::|||
Db       1295 CCAGTCGATGGGGCGTGGACGGAGTGGAGCAAGTGGTCAGCCTGCAGCACTGAGTGTGCC 1354

Qy        317 HisTrpArgSerArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGln 336
                 ||||||||||||||||||||||       |||   |||:::||||||||    :::|||
Db       1355 CACTGGCGTAGCCGCGAGTGCATGGCGCCCCCACCCCAGAACGGAGGCCGTGACTGCAGC 1414

Qy        337 GlyThrAspLeuAspThrArgAsnCysThrSerAspLeuCysValHisSer--------- 353
                 |||||   ||||||::::::|||||||||      ||||||:::  :::
Db       1415 GGGACGCTGCTCGACTCTAAGAACTGCACAGATGGGCTGTGCATGCAAAATAAGAAAACT 1474

Qy        354 ---------------------------AlaSerGlyProGluAspValAlaLeuTyr 363
                                         |||||||||     |||   ||||||||
Db       1475 CTAAGCGACCCCAACAGCCACCTGCTGGAGGCCTCAGGG------GATGCGGCGCTGTAT 1528

Qy        364 ValGlyLeu---IleAlaValAlaValCysLeuValLeuLeuLeuLeuValLeuIleLeu 382
                 ||||||    :::|||:::    |||    :::    :::|||:::    |||    ::::::
Db       1529 GCGGGGCTCGTGGTGGCCATCTTCGTGGTCGTGGCAATCCTCATGGCGGTGGGGGTGGTG 1588

Qy        383 ValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSer---IleLeu 401
                 ||||||   |||             |||:::|||:::   |||||||||           |||
Db       1589 GTGTACCGCCGCAACTGCCGTGACTTCGACACAGACATCACTGACTCATCTGCTGCCCTG 1648

Qy        402 ThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeuLeu 421
                 |||   ||||||   |||||||:::   |||    ::::::     ||||||     ||||||
Db       1649 ACTGGTGGTTTCCACCCCGTCAACTTTAAGACGGCAAGGCCCAGCAACCCGCAGCTCCTA 1708

Qy        422 ------ThrIleGlnProAspLeuSerThrThrThr---ThrTyrGlnGlySerLeuCys 438
                    ::::::    |||||||||:::    :::          |||:::|||       :::
Db       1709 CACCCCTCTGTGCCTCCTGACCTGACAGCCAGCGCCGGCATCTACCGCGGACCCGTGTAT 1768

Qy        439 ProArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerPro 458
                       ||||||        :::    |||        :::||||||       ||||||     |||
Db       1769 GCCCTGCAGGACTCC---ACCGACAAAATCCCCATGACCAACTCTCCTCTGCTGGACCCC 1825

Qy        459 LeuGlyGlyGlyArgHisThrLeuHisHisSerSerProThr------------------ 472
                 |||          :::       ::::::   ||||||    |||
Db       1826 TTACCCAGCCTTAAGGTCAAGGTCTACAGCTCCAGCACCACGGGCTCTGGGCCAGGCCTG 1885

Qy        473 SerGluAlaGluGluPheValSerArgLeuSerThrGlnAsnTyr--------------- 487
                 ::::::      :::    :::              |||
Db       1886 GCAGATGGGGCTGACCTGCTGGGGGTCTTGCCGCCTGGCACATACCCTAGCGATTTCGCC 1945

Qy        488 --------------------PheArgSer----------------------------- 490
                                     |||||
Db       1946 CGGGACACCCACTTCCTGCACCTGCGCAGCGCCAGCCTCGGTTCCCAGCAGCTCTTGGGC 2005

Qy        491 LeuProArgGlyThrSerAsnMetThrTyrGlyThrPheAsnPheLeuGlyGlyArgLeu 510
                 |||||||||          :::          |||||||||     ||||||||||||||
Db       2006 CTGCCCCGAGACCCAGGGAGCAGCGTCAGCGGCACCTTTGGCTGCCTGGGTGGAGGCTC 2065

Qy        511 MetIleProAsnThrGlyIleSerLeuLeuIleProProAspAlaIleProArgGlyLys 530
                 ||||||    ||||||:::||||||||:::|||         |||||||||:::||||||
Db       2066 AGCATCCCCGGCACAGGGGTCAGCTTGCTGGTGCCCAATGGAGCCATTCCCCAGGGCAAG 2125

Qy        531 IleTyrGluIleTyrLeuThrLeuHisLysProGluAspValArgLeuProLeuAla--- 549
                 ||||||:::||||||       ::::::|||   |||           |||||||||:::
Db       2126 TTCTACGAGATGTATCTACTCATCAACAAGGCAGAAAGTACC---CTCCCGCTTTCAGAA 2182

Qy        550 GlyCysGlnThrLeuLeuSerProIleValSerCysGlyProProGlyValLeuLeuThr 569
                 |||      ||||||:::|||||||||     |||:::|||||||||     |||:::||||||
Db       2183 GGGACCCAGACAGTATTGAGCCCCTCGGTGACCTGTGGACCCACAGGCCTCCTGCTGTGC 2242

Qy        570 ArgProValIleLeuAlaMetAspHisCysGlyGluProSerProAspSerTrpSerLeu 589
                 ||||||||||||||    |||    ||||||    |||   |||              |||
Db       2243 CGCCCCGTCATCCTCACCATGCCCCACTGTGCCGAAGTCAGTGCCCGTGACTGGATCTTT 2302

Qy        590 ArgLeuLysLysGlnSerCysGluGlySerTrpGluGlnAspValLeuHisLeuGlyGlu 609
                 :::||||||    |||:::    :::|||    ||||||    :::|||:::     |||    |||
Db       2303 CAGCTCAAGACCCAGGCCCACCAGGGCCACTGGGAG---GAGGTGGTGACCCTGGATGAG 2359

Qy        610 GluAlaProSerHisLeuTyrTyrCysGlnLeuGluAlaSerAlaCysTyrValPheThr 629
                 |||         :::           ||||||||||||||        ||||||:::::
Db       2360 GAGACCCTGAACACACCCTGCTACTGCCAGCTGGAGCCCAGGGCCTGTCACATCCTGCTG 2419

Qy        630 GluGlnLeuGlyArgPheAlaLeuValGlyGluAlaLeuSerValAlaAlaAlaLysArg 649
                 :::|||||||||    :::          ||||||:::   |||    :::|||   ||||||
Db       2420 GACCAGCTGGGCACCTACGTGTTCACGGGCGAGTCCTATTCCCGCTCAGCAGTCAAGCGG 2479

Qy        650 LeuLysLeuLeuLeuPheAlaProValAlaCysThrSerLeuGluTyrAsnIleArgVal 669
                 |||:::|||    :::|||||||||      |||||||||||||||||:::::::||||||
Db       2480 CTCCAGCTGGCCGTCTTCGCCCCCGCCCTCTGCACCTCCCTGGAGTACAGCCTCCGGGTC 2539

Qy        670 TyrCysLeuHisAspThrHisAspAlaLeuLysGluValValGlnLeuGluLysGlnLeu 689
                 |||||||||    ||||||         ||||||||||||||||||:::::::||||||:::   |||
Db       2540 TACTGCCTGGAGGACACGCCTGTAGCACTGAAGGAGGTGCTGGAGCTGGAGCGGACTCTG 2599

Qy        690 GlyGlyGlnLeuIleGlnGluProArgValLeuHisPheLysAspSerTyrHisAsnLeu 709
                 ||||||   |||::::::|||||||||:::    |||    |||||||||||||||||||||||||
Db       2600 GGCGGATACTTGGTGGAGGAGCCGAAACCGCTAATGTTCAAGGACAGTTACCACAACCTG 2659

Qy        710 ArgLeuSerIleHisAspValProSerSerLeuTrpLysSerLysLeuLeuValSerTyr 729
                 ||||||||:::||||||:::|||    :::      |||:::||||||||||||         |||
Db       2660 CGCCTCTCCCTCCATGACCTCCCCCATGCCCATTGGAGGAGCAAGCTGCTGGCCAAATAC 2719

Qy        730 GlnGluIleProPheTyrHisIleTrpAsnGlyThrGlnArgTyrLeuHisCysThrPhe 749
                 ||||||||||||||||||||||||:::|||:::|||:::      |||||||||||||||
Db       2720 CAGGAGATCCCCTTCTATCACATTTGGAGTGGCAGCCAGAAGGCCCTCCACTGCACTTTC 2779

Qy        750 ThrLeuGluArgValSerProSerThrSerAspLeuAlaCysLysLeuTrpValTrpGln 769
                 ||||||||||||   |||    ::::::::::::|||    ||||||:::      |||    |||
Db       2780 ACCCTGGAGAGGCACAGCTTGGCCTCCACAGAGCTCACCTGCAAGATCTGCGTGCGGCAA 2839

Qy        770 ValGluGlyAspGlyGlnSerPheSerIleAsnPheAsnIleThrLysAspThrArgPhe 789
                 |||||||||:::||||||   ||||    ::::::          :::         :::|||
Db       2840 GTGGAAGGGGAGGGCCAGATATTCCAGCTGCATACCACTCTGGCA---GAGACACCTGCT 2896

Qy        790 AlaGluLeuLeuAlaLeuGluSerGluAlaGly------ValProAlaLeuValGlyPro 807
                       |||      |||   |||      |||          |||          :::||||||
Db       2897 GGCTCCCTGGACACTCTCTGCTCTGCCCCTGGCAGCACTGTCACCACCCAGCTGGGACCT 2956

Qy        808 SerAlaPheLysIleProPheLeuIleArgGlnLysIleIleSerSerLeuAspProPro 827
                    |||||||||||||||     |||||||||||||||     :::|||||||||    |||
Db       2957 TATGCCTTCAAGATCCCACTGTCCATCCGCCAGAAGATATGCAACAGCCTAGATGCCCCC 3016

Qy        828 CysArgArgGlyAlaAspTrpArgThrLeuAlaGlnLysLeuHisLeuAspSerHisLeu 847
                       ||||||    |||||||||   |||||||||||||||     :::|||    :::|||
Db       3017 AACTCACGGGGCAATGACTGGCGGATGTTAGCACAGAAGCTCTCTATGGACCGGTACCTG 3076

Qy        848 SerPhePheAlaSerLysProSerProThrAlaMetIleLeuAsnLeuTrpGluAlaArg 867
                 ::::::||||||:::|||    |||||||||    :::||||||:::||||||||||||
Db       3077 AATTACTTTGCCACCAAAGCGAGCCCCACGGGTGTGATCCTGGACCTCTGGGAAGCTCTG 3136

Qy        868 HisPheProAsnGlyAsnLeuSerGlnLeuAlaAlaAlaValAlaGlyLeuGlyGlnPro 887
                          :::|||:::|||:::    ||||||:::|||:::          :::|||:::
Db       3137 CAGCAGGACGATGGGGACCTCAACAGCCTGGCGAGTGCCTTGGAGGAGATGGGCAAGAGT 3196

Qy        888 AspAlaGlyLeuPheThrValSerGluAlaGluCys 899
                 :::       :::          ::::::   :::|||
Db       3197 GAGATGCTGGTGGCTGTGGCCACCGACGGGGACTGC 3232
```

Search completed: October  6, 2004, 18:38:43
Job time : 971 secs

GenCore version 5.1.6


```
OM protein - nucleic search, using frame_plus_p2n model

Run on:         October  6, 2004, 18:19:39 ; Search time 8838 Seconds
                                             (without alignments)
                                             4408.847 Million cell updates/sec

Title:          US-09-970-944-2
Perfect score:  4787
Sequence:       1 MAVRPGLWPALLGIVLAAWL..........AVAGLGQPDAGLFTVSEAEC 899

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters:      6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09970944/runat_05102004_112010_1626/app_query.fasta_1.1095
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09970944_@CGN_1_1_5932_@runat_05102004_112010_1626 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      GenEmbl:*
                1:  gb_ba:*
                2:  gb_htg:*
                3:  gb_in:*
                4:  gb_om:*
                5:  gb_ov:*
                6:  gb_pat:*
                7:  gb_ph:*
                8:  gb_pl:*
                9:  gb_pr:*
                10:  gb_ro:*
                11:  gb_sts:*
                12:  gb_sy:*
                13:  gb_un:*
                14:  gb_vi:*
                15:  em_ba:*
                16:  em_fun:*
                17:  em_hum:*
                18:  em_in:*
                19:  em_mu:*
                20:  em_om:*
                21:  em_or:*
                22:  em_ov:*
                23:  em_pat:*
                24:  em_ph:*
                25:  em_pl:*
                26:  em_ro:*
                27:  em_sts:*
                28:  em_un:*
                29:  em_vi:*
                30:  em_htg_hum:*
                31:  em_htg_inv:*
                32:  em_htg_other:*
                33:  em_htg_mus:*
                34:  em_htg_pln:*
                35:  em_htg_rod:*
                36:  em_htg_mam:*
                37:  em_htg_vrt:*
                38:  em_sy:*
                39:  em_htgo_hum:*
                40:  em_htgo_mus:*
                41:  em_htgo_other:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| | Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|---|
| | 1 | 4787 | 100.0 | 2881 | 6 | AX527916 | AX527916 Sequence |
| | 2 | 4698.5 | 98.2 | 2752 | 6 | AX449572 | AX449572 Sequence |
| | 3 | 4688.5 | 97.9 | 2697 | 6 | AX451652 | AX451652 Sequence |
| | 4 | 4592.5 | 95.9 | 3992 | 10 | MMU487852 | AJ487852 Mus muscu |
| | 5 | 4545.5 | 95.0 | 2697 | 6 | AX268596 | AX268596 Sequence |
| | 6 | 4545.5 | 95.0 | 2697 | 10 | RNU87305 | U87305 Rattus norv |
| | 7 | 4545.5 | 95.0 | 3014 | 6 | BD057524 | BD057524 Netrin re |
| | 8 | 4379.5 | 91.5 | 3580 | 6 | AX367094 | AX367094 Sequence |
| | 9 | 4278.5 | 89.4 | 3844 | 10 | BC058084 | BC058084 Mus muscu |
| | 10 | 3024.5 | 63.2 | 4294 | 10 | AK122575 | AK122575 Mus muscu |
| | 11 | 2819 | 58.9 | 2688 | 9 | BC009333 | BC009333 Homo sapi |
| | 12 | 2705.5 | 56.5 | 9299 | 10 | MMU72634 | U72634 Mus musculu |
| | 13 | 2694 | 56.3 | 1787 | 6 | BD057525 | BD057525 Netrin re |
| | 14 | 2688.5 | 56.2 | 9328 | 10 | AB118026 | AB118026 Rattus no |
| c | 15 | 2681 | 56.0 | 200968 | 2 | AC123700 | AC123700 Mus muscu |
| | 16 | 2679.5 | 56.0 | 2962 | 5 | AY187310 | AY187310 Gallus ga |
| | 17 | 2673.5 | 55.8 | 3646 | 9 | AF055634 | AF055634 Homo sapi |
| | 18 | 2570 | 53.7 | 2832 | 5 | AY099459 | AY099459 Xenopus l |
| c | 19 | 2513.5 | 52.5 | 166534 | 2 | AC034209 | AC034209 Homo sapi |
| | 20 | 2505.5 | 52.3 | 3672 | 10 | BC057560 | BC057560 Mus muscu |
| | 21 | 2505.5 | 52.3 | 3672 | 10 | BC048162 | BC048162 Mus muscu |
| | 22 | 2499 | 52.2 | 3788 | 10 | MMU487853 | AJ487853 Mus muscu |
| | 23 | 2497 | 52.2 | 2838 | 10 | RNU87306 | U87306 Rattus norv |
| | 24 | 2490.5 | 52.0 | 2995 | 6 | AX497288 | AX497288 Sequence |
| | 25 | 2486.5 | 51.9 | 3770 | 9 | AY126437 | AY126437 Homo sapi |
| | 26 | 2484 | 51.9 | 2860 | 6 | AX686445 | AX686445 Sequence |
| | 27 | 2484 | 51.9 | 2860 | 6 | AX686447 | AX686447 Sequence |
| | 28 | 2479 | 51.8 | 3884 | 6 | AX464012 | AX464012 Sequence |
| | 29 | 2479 | 51.8 | 3884 | 9 | AY358351 | AY358351 Homo sapi |
| | 30 | 2479 | 51.8 | 3933 | 9 | AK128132 | AK128132 Homo sapi |
| | 31 | 2479 | 51.8 | 3935 | 9 | AB096256 | AB096256 Homo sapi |
| | 32 | 2472 | 51.6 | 2895 | 6 | AX512281 | AX512281 Sequence |
| | 33 | 2463 | 51.5 | 2831 | 6 | BD057526 | BD057526 Netrin re |
| c | 34 | 2444.5 | 51.1 | 168168 | 2 | AC012283 | AC012283 Homo sapi |
| | 35 | 2381 | 49.7 | 121282 | 9 | AC027318 | AC027318 Homo sapi |
| | 36 | 2276.5 | 47.6 | 9700 | 6 | AX054976 | AX054976 Sequence |
| c | 37 | 2194.5 | 45.8 | 231407 | 2 | AC139592 | AC139592 Rattus no |
| | 38 | 2148 | 44.9 | 2868 | 6 | AX800719 | AX800719 Sequence |
| | 39 | 2146 | 44.8 | 3252 | 9 | AY358147 | AY358147 Homo sapi |
| | 40 | 2144.5 | 44.8 | 3925 | 10 | MMU487854 | AJ487854 Mus muscu |
| | 41 | 2140.5 | 44.7 | 2661 | 6 | AX800717 | AX800717 Sequence |
| | 42 | 2136 | 44.6 | 7002 | 6 | BD162193 | BD162193 Novel unc |
| | 43 | 2136 | 44.6 | 7002 | 9 | AB055056 | AB055056 Homo sapi |
| | 44 | 1646 | 34.4 | 6495 | 9 | AB075856 | AB075856 Homo sapi |
| c | 45 | 1646 | 34.4 | 6522 | 6 | BD186447 | BD186447 Novel gen |

ALIGNMENTS

RESULT 1

```
AX527916
LOCUS       AX527916                2881 bp    DNA     linear   PAT 21-NOV-2002
DEFINITION  Sequence 1 from Patent WO0229038.
ACCESSION   AX527916
VERSION     AX527916.1  GI:25172359
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Herrmann,J.L., Rastelli,L. and Shimkets,R.A.
  TITLE     Novel proteins and nucleic acids encoding same and antibodies
            directed against these proteins
  JOURNAL   Patent: WO 0229038-A 1 11-APR-2002;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1..2881
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              6.13e-295     Length:          2881
Score:                  4787.00       Matches:         899
Percent Similarity:     100.00%       Conservative:    0
Best Local Similarity:  100.00%       Mismatches:      0
Query Match:            100.00%       Indels:          0
DB:                     6             Gaps:            0

US-09-970-944-2 (1-899) x AX527916 (1-2881)

Qy          1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         87 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 146

Qy         21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 206

Qy         41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 266

Qy         61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 326

Qy         81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCG 386

Qy        101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 446

Qy        121 GluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 506

Qy        141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 566

Qy        161 SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 626

Qy        181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 686

Qy        201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 746

Qy        221 CysValAlaLysAsnIleValAlaArgArgArgSerAlaSerAlaAlaValIleValTyr 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        747 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 806

Qy        241 ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        807 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 866

Qy        261 GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        867 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 926

Qy        281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        927 TGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGC 986

Qy        301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        987 AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC 1046

Qy        321 ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1047 CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG 1106

Qy        341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1107 GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG 1166

Qy        361 AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuLeuValLeu 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1167 GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC 1226

Qy        381 IleLeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIle 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1227 ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1286

Qy        401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1287 CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1346

Qy        421 LeuThrIleGlnProAspLeuSerThrThrThrThrTyrGlnGlySerLeuCysProArg 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1347 CTCACCATCCAGCCGGACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGG 1406

Qy        441 GlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeuGly 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1407 CAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGT 1466

Qy        461 GlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheValSer 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1467 GGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCC 1526

Qy        481 ArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThrTyr 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1527 CGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT 1586

Qy        501 GlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeuLeu 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1587 GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTC 1646

Qy        521 IleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHisLys 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1647 ATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG 1706

Qy        541 ProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleValSer 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1707 CCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC 1766

Qy        561 CysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGly 580
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1767 TGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG 1826

Qy        581 GluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySerTrp 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1827 GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG 1886

Qy        601 GluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGlnLeu 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1887 GAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTG 1946

Qy        621 GluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGlu 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1947 GAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAG 2006

Qy        641 AlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAlaCys 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2007 GCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGC 2066

Qy        661 ThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuLys 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2067 ACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAG 2126

Qy        681 GluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgValLeu 700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2127 GAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTG 2186

Qy        701 HisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSerLeu 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2187 CACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTG 2246

Qy        721 TrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsnGly 740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2247 TGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGC 2306

Qy        741 ThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSerAsp 760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2307 ACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGAC 2366

Qy        761 LeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIleAsn 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2367 CTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAAC 2426

Qy        781 PheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAlaGly 800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2427 TTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGG 2486

Qy        801 ValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLysIle 820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2487 GTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATA 2546

Qy        821 IleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGlnLys 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2547 ATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAA 2606

Qy        841 LeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMetIle 860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2607 CTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATC 2666

Qy        861 LeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAlaAla 880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2667 CTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCA 2726

Qy        881 ValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2727 GTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2783
```

RESULT 2
AX449572

```
LOCUS       AX449572                2752 bp    DNA     linear   PAT 03-JUL-2002
DEFINITION  Sequence 1 from Patent WO0210216.
ACCESSION   AX449572
VERSION     AX449572.1  GI:21698195
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Padigaru,M., Mezes,P., Mishra,V., Burgess,C., Casman,S.,
            Grosse,W.M., Alsobrook,J.P., Lepley,D.M., Gerlach,V.L.,
            Macdougall,J.R. and Smithson,G.
  TITLE     Proteins and nucleic acids encoding same
  JOURNAL   Patent: WO 0210216-A 1 07-FEB-2002;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1..2752
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              2.45e-289      Length:         2752
Score:                  4698.50        Matches:        888
Percent Similarity:     98.89%         Conservative:   2
Best Local Similarity:  98.67%         Mismatches:     7
Query Match:            98.15%         Indels:         3
DB:                     6              Gaps:           3

US-09-970-944-2 (1-899) x AX449572 (1-2752)

Qy          1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         46 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 105

Qy         21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        106 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 165

Qy         41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        166 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 225

Qy         61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        226 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 285

Qy         81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100
              |||||||||||||||||||||||||||||||||||||||||||||||||||     |||
Db        286 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 345

Qy        101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        346 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 405

Qy        121 GluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        406 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 465

Qy        141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        466 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 525

Qy        161 SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        526 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 585

Qy        181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        586 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 645

Qy        201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        646 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 705
```

```
Qy     221 CysValAlaLysAsnIleValAlaArgArgArgSerAlaSerAlaAlaValIleValTyr 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     706 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 765

Qy     241 ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg 260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     766 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 825

Qy     261 GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe 280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     826 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 885

Qy     281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300
           ||||||||||||||||||   :::|||      :::   |||       ||||||||
Db     886 TGTGAGGGGCAGAATGTC---CAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGC 942

Qy     301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     943 AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC 1002

Qy     321 ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeu 340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1003 CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG 1062

Qy     341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1063 GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG 1122

Qy     361 AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuLeuValLeu 380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1123 GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC 1182

Qy     381 IleLeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIle 400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1183 ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1242

Qy     401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1243 CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1302

Qy     421 LeuThrIleGlnProAspLeuSer---ThrThrThrThrTyrGlnGlySerLeuCysPro 439
           |||||||||||||||||||||      |||||||||||||||||||||||||||||||||
Db    1303 CTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCC 1362

Qy     440 ArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu 459
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1363 CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG 1422

Qy     460 GlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheVal 479
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1423 GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC 1482

Qy     480 SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr 499
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1483 TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC 1542

Qy     500 TyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeu 519
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1543 TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC 1602

Qy     520 LeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHis 539
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1603 CTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC 1662

Qy     540 LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal 559
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1663 AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT 1722

Qy     560 SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys 579
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1723 AGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT 1782

Qy     580 GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySer 599
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1783 GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC 1842

Qy     600 TrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGln 619
           ||||||   ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1843 TGGGAG---GATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG 1899

Qy     620 LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly 639
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1900 CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA 1959

Qy     640 GluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAla 659
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1960 GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC 2019

Qy     660 CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu 679
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2020 TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC 2079

Qy     680 LysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgVal 699
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2080 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTC 2139

Qy     700 LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2140 CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCC 2199

Qy     720 LeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn 739
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2200 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2259

Qy     740 GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2260 GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2319

Qy     760 AspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIle 779
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2320 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2379

Qy     780 AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAla 799
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2380 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2439

Qy     800 GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2440 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2499

Qy     820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2500 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2559

Qy     840 LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2560 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG 2619

Qy     860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2620 ATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2679

Qy     880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2680 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2739
```

```
RESULT 3
AX451652
LOCUS       AX451652                2697 bp    DNA     linear   PAT 03-JUL-2002
DEFINITION  Sequence 1 from Patent WO0233080.
ACCESSION   AX451652
VERSION     AX451652.1  GI:21698587
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Koehler,R.H.
  TITLE     Regulation of human netrin binding membrane receptor unc5h-1
  JOURNAL   Patent: WO 0233080-A 1 25-APR-2002;
            Bayer Aktiengesellschaft (DE)
FEATURES             Location/Qualifiers
     source          1..2697
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                1.03e-288      Length:          2697
Score:                    4688.50        Matches:         886
Percent Similarity:       98.78%         Conservative:    3
Best Local Similarity:    98.44%         Mismatches:      8
Query Match:              97.94%         Indels:          3
DB:                       6              Gaps:            3

US-09-970-944-2 (1-899) x AX451652 (1-2697)

Qy          1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 60

Qy         21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120

Qy         41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 180

Qy         61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 240

Qy         81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100
              |||||||||||||||||||||||||||||||||||||||||||||||||||||    |||
Db        241 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 300

Qy        101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 360

Qy        121 GluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 420

Qy        141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160
              |||||||||||||||     ||||||||||||||||||||||||||||||||||||||||
Db        421 TACATCCGCATAGCCTATTTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 480

Qy        161 SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 540

Qy        181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 600

Qy        201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 660

Qy        221 CysValAlaLysAsnIleValAlaArgArgArgSerAlaSerAlaAlaValIleValTyr 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 720

Qy        241 ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 780

Qy        261 GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 840

Qy        281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300
              ||||||||||||||||||         :::|||     :::    |||     ||||||||
Db        841 TGTGAGGGGCAGAATGTC---CAGAAAACAGCCTGCGCCACCCTGTGCCCAGTGGACGGC 897

Qy        301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        898 AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC 957

Qy        321 ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        958 CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG 1017

Qy        341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360
              ||||||||||||||||||||||||||||||:::|||||||||||||||||||||||||||
Db       1018 GACACCCGCAACTGTACCAGTGACCTCTGTGTACACACTGCTTCTGGCCCTGAGGACGTG 1077

Qy        361 AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuValLeu 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1078 GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC 1137

Qy        381 IleLeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIle 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1138 ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1197

Qy        401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1198 CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1257

Qy        421 LeuThrIleGlnProAspLeuSer---ThrThrThrThrTyrGlnGlySerLeuCysPro 439
              |||||||||||||||||||||||   ||||||||||||||||||||||||||||||||||
Db       1258 CTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCC 1317

Qy        440 ArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu 459
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1318 CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG 1377

Qy        460 GlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheVal 479
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1378 GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC 1437

Qy        480 SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr 499
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1438 TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC 1497

Qy        500 TyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeu 519
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1498 TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGAATCAGCCTC 1557

Qy        520 LeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHis 539
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1558 CTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC 1617

Qy        540 LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal 559
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1618 AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT 1677

Qy        560 SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys 579
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1678 AGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT 1737

Qy        580 GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySer 599
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1738 GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC 1797

Qy        600 TrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGln 619
              ||||||   |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1798 TGGGAG---GATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG 1854
```

```
Qy        620 LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly 639
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1855 CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA 1914

Qy        640 GluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAla 659
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1915 GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC 1974

Qy        660 CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu 679
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1975 TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC 2034

Qy        680 LysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgVal 699
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2035 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTC 2094

Qy        700 LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2095 CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCC 2154

Qy        720 LeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn 739
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2155 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2214

Qy        740 GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2215 GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2274

Qy        760 AspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIle 779
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2275 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2334

Qy        780 AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAla 799
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2335 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2394

Qy        800 GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2395 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2454

Qy        820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2455 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2514

Qy        840 LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2515 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG 2574

Qy        860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2575 ATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2634

Qy        880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2635 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2694
```

```
RESULT 4
MMU487852
LOCUS       MMU487852             3992 bp    mRNA    linear   ROD 24-SEP-2002
DEFINITION  Mus musculus mRNA for netrin receptor Unc5h1 (Unc5h1 gene).
ACCESSION   AJ487852
VERSION     AJ487852.1  GI:22035783
KEYWORDS    netrin receptor Unc5h1; Unc5h1 gene.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Engelkamp,D.
  TITLE     Cloning of three mouse Unc5 genes and their expression patterns at
            mid-gestation
  JOURNAL   Mech. Dev. 118 (1-2), 191-197 (2002)
   MEDLINE  22239710
    PUBMED  12351186
REFERENCE   2  (bases 1 to 3992)
  AUTHORS   Engelkamp,D.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-MAY-2002) Neuroanatomy, Max Planck Institute for
            Brain Research, Deutschordenstrasse 46, Frankfurt 60528, GERMANY
FEATURES             Location/Qualifiers
     source          1..3992
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
     gene            1..3992
                     /gene="Unc5h1"
     CDS             232..2928
                     /gene="Unc5h1"
                     /codon_start=1
                     /product="netrin receptor Unc5h1"
                     /protein_id="CAD32250.1"
                     /db_xref="GI:22035784"
                     /db_xref="GOA:Q8K1S4"
                     /db_xref="SPTREMBL:Q8K1S4"
                     /translation="MAVRPGLWPALLGIVLTAWLRGSGAQQSATVANPVPGANPDLLP
                     HFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTM
                     EVRINVSRQQVEKVFGLEEYWCQCVAWSSSGTTKSQKAYIRIAYLRKNFEQEPLAKEV
                     SLEQGIVLPCRPPEGIPPAEVEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTAN
                     YTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLN
                     GGAFCEGQNVQKTACATLCPVDGSWSPWSKWSACGLDCTHWRSRECSDPAPRNGGEEC
                     RGADLDTRNCTSDLCLHTSSGPEDVALYIGLVAVAVCLILLLLVLVLIYCRKKEGLDS
                     DVADSSILTSGFQPVSIKPSKADNPHLLTIQPDLSTTTTTYQGSLCPRQDGPSPKFQL
                     SNGHLLSPLGSGRHTLHHSSPTSEAEDFVSRLSTQNYFRSLPRGTSNMAYGTFNFLGG
                     RLMIPNTGISLLIPPDAIPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGV
                     LLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGSWEDVLHLGEESPSHLYYCQLEAGAC
                     YVFTEQLGRFALVGEALSVAATKRLRLLLFAPVACTSLEYNIRVYCLHDTHDALKEVV
                     QLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWKSKLLVSYQEIPFYHIWNGT
                     QQYLHCTFTLERVNASTSDLACKVWVWQVEGDGQSFNINFNITKDTRFAEMLALESEG
                     GVPALVGPSAFKIPFLIRQKIITSLDPPCSRGADWRTLAQKLHLDSHLSFFASKPSPT
                     AMILNLWEARHFPNGNLGQLAAAVAGLGQPDAGLFTVSEAEC"
ORIGIN

Alignment Scores:
Pred. No.:              2.19e-282      Length:          3992
Score:                  4592.50        Matches:         858
Percent Similarity:     97.67%         Conservative:    21
Best Local Similarity:  95.33%         Mismatches:      18
Query Match:            95.94%         Indels:          3
DB:                     10             Gaps:            3

US-09-970-944-2 (1-899) x MMU487852 (1-3992)
```

```
Qy          1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||    ||||||||||
Db        232 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCACTGCCTGGCTT 291

Qy         21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 CGTGGTTCGGGTGCCCAGCAGAGTGCCACAGTGGCCAACCCAGTGCCTGGTGCCAACCCA 351

Qy         41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 GACCTGCTGCCCCACTTCCTGGTAGAGCCGGAGGACGTGTACATTGTCAAGAACAAGCCC 411

Qy         61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 GTGCTGCTGGTGTGCAAGGCTGTGCCCGCCACCCAGATCTTCTTCAAGTGCAACGGGGAA 471

Qy         81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||    |||
Db        472 TGGGTTCGCCAGGTCGATCACGTCATCGAACGCAGCACTGACGGCAGCAGCGGATTGCCA 531

Qy        101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        532 ACCATGGAGGTCCGGATCAACGTATCAAGGCAGCAGGTCGAGAAAGTGTTTGGGCTGGAG 591
```

```
Qy    121 GluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAla 140
Db    592 GAGTACTGGTGCCAGTGTGTGGCATGGAGCTCCTCAGGAACCACCAAAAGCCAGAAGGCC 651

Qy    141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160
Db    652 TACATCCGGATTGCCTATTTGCGCAAGAACTTTGAGCAGGAGCCGCTGGCCAAGGAAGTG 711

Qy    161 SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu 180
Db    712 TCACTGGAGCAAGGCATTGTGCTACCTTGTCGCCCCCCGGAAGGAATCCCCCCAGCTGAG 771

Qy    181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200
Db    772 GTGGAGTGGCTCCGAAATGAGGACCTCGTGGACCCCTCCCTCGACCCCAATGTGTACATC 831

Qy    201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220
Db    832 ACACGGGAGCACAGCCTAGTCGTGCGGCAGGCCCGCCTGGCCGACACTGCCAACTACACC 891

Qy    221 CysValAlaLysAsnIleValAlaArgArgArgSerAlaSerAlaAlaValIleValTyr 240
Db    892 TGCGTGGCCAAGAACATCGTGGCCCGTCGCCGAAGCGCCTCTGCGGCCGTCATTGTTTAT 951

Qy    241 ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg 260
Db    952 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGTGCCAGCTGTGGGCGT 1011

Qy    261 GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe 280
Db   1012 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCACCTCTCAACGGGGGCGCCTTC 1071

Qy    281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300
Db   1072 TGTGAGGGGCAGAATGTC---CAGAAAACAGCCTGCGCCACTCTGTGCCCAGTGGATGGG 1128

Qy    301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320
Db   1129 AGCTGGAGCCCATGGAGTAAGTGGTCAGCCTGCGGGCTTGACTGCACCCACTGGCGGAGC 1188

Qy    321 ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeu 340
Db   1189 CGGGAGTGCTCCGACCCAGCGCCCCGCAACGGAGGTGAGGAGTGCCGGGGTGCTGACCTG 1248

Qy    341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360
Db   1249 GACACCCGCAACTGTACCAGTGACCTCTGCCTGCACACCTCTTCCGGCCCCGAGGACGTG 1308

Qy    361 AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuLeuValLeu 380
Db   1309 GCTCTCTACATCGGCCTCGTCGCCGTGGCCGTGTGCCTCATCTTGCTGCTGCTGGTCCTC 1368

Qy    381 IleLeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIle 400
Db   1369 GTCCTCATCTACTGCCGCAAGAAGGAAGGACTGGACTCAGACGTGGCTGACTCATCCATC 1428

Qy    401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420
Db   1429 CTTACCTCAGGCTTCCAGCCTGTCAGCATCAAGCCCAGCAAAGCAGACAATCCCCATCTG 1488

Qy    421 LeuThrIleGlnProAspLeuSer---ThrThrThrThrTyrGlnGlySerLeuCysPro 439
Db   1489 CTCACCATCCAACCGGACCTCAGCACCACCACGACCACCTACCAGGGCAGCCTGTGTCCC 1548

Qy    440 ArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu 459
Db   1549 CGGCAGGATGGACCCAGCCCCAAGTTCCAGCTCTCTAATGGTCACCTGCTCAGCCCACTG 1608

Qy    460 GlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheVal 479
Db   1609 GGCAGTGGCCGCCATACGCTGCACCACAGCTCCCCCACCTCTGAGGCTGAGGACTTCGTC 1668

Qy    480 SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr 499
Db   1669 TCCCGCCTCTCCACCCAAAACTACTTTCGTTCTCTGCCCCGCGGTACCAGCAACATGGCC 1728

Qy    500 TyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeu 519
Db   1729 TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAACACAGGAATCAGCCTC 1788

Qy    520 LeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHis 539
Db   1789 CTCATACCCCCGGACGCCATCCCCCGAGGAAAGATCTACGAGATCTACCTCACTCTGCAC 1848

Qy    540 LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal 559
Db   1849 AAGCCAGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCTATCGTT 1908

Qy    560 SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys 579
Db   1909 AGCTGTGGGCCCCCAGGAGTCCTGCTCACCCGGCCAGTCATCCTTGCCATGGACCACTGC 1968

Qy    580 GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySer 599
Db   1969 GGGGAGCCCAGTCCCGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCCTGTGAGGGCAGC 2028

Qy    600 TrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGln 619
Db   2029 TGGGAG---GACGTGCTGCACCTTGGTGAGGAGTCGCCCTCTCATCTCTACTACTGCCAG 2085

Qy    620 LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly 639
Db   2086 CTGGAGGCCGGGGCCTGCTATGTCTTCACCGAGCAGCTAGGCCGCTTTGCCCTGGTGGGA 2145

Qy    640 GluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAla 659
Db   2146 GAGGCCCTCAGCGTGGCTGCCACCAAGCGCCTCAGGCTCCTTCTGTTTGCCCCTGTGGCC 2205

Qy    660 CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu 679
Db   2206 TGTACGTCCCTCGAGTACAACATCCGAGTGTACTGCCTGCACGACACCCACGATGCTCTC 2265

Qy    680 LysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgVal 699
Db   2266 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGTGGACAGCTGATCCAGGAGCCCCGTGTC 2325

Qy    700 LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719
Db   2326 CTGCACTTCAAAGACAGTTACCACAACCTACGTCTGTCCATCCACGACGTGCCCAGCTCC 2385

Qy    720 LeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn 739
Db   2386 CTGTGGAAGAGCAAGCTCCTTGTCAGCTACCAGGAGATCCCTTTTTACCACATCTGGAAT 2445

Qy    740 GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759
Db   2446 GGCACTCAGCAGTATCTGCACTGCACCTTCACCCTGGAGCGCGTCAATGCCAGCACCAGC 2505

Qy    760 AspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIle 779
Db   2506 GACCTGGCCTGCAAGGTGTGGGTGTGGCAGGTGGAGGGAGATGGACAGAGCTTCAACATC 2565

Qy    780 AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAla 799
Db   2566 AACTTTAACATCACTAAGGACACGAGGTTTGCTGAAATGCTGGCTCTGGAGAGTGAAGGG 2625

Qy    800 GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
Db   2626 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAAAAG 2685

Qy    820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839
Db   2686 ATCATTACCAGCCTGGACCCACCCTGCAGCCGGGGCGCCGACTGGCGAACTCTAGCCCAG 2745

Qy    840 LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859
```

```
Db       2746 AAACTTCACCTGGACAGCCATCTTAGCTTCTTTGCCTCCAAGCCCAGCCCTACAGCCATG 2805

Qy        860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879

Db       2806 ATCCTCAACCTATGGAGGCGCGGCACTTCCCCAACGGCAACCTCGGCCAGCTGGCCGCA 2865

Qy        880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899

Db       2866 GCTGTGGCCGGACTGGGCCAGCCAGATGCTGGCCTCTTCACCGTGTCAGAGGCCGAGTGC 2925

RESULT 5
AX268596
LOCUS       AX268596                 2697 bp    DNA     linear   PAT 29-OCT-2001
DEFINITION  Sequence 15 from Patent WO0175440.
ACCESSION   AX268596
VERSION     AX268596.1  GI:16541710
KEYWORDS    .
SOURCE      Rattus sp.
  ORGANISM  Rattus sp.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1
  AUTHORS   Cochran,S.W., Paterson,G.Y., Ohashi,Y.W., Morris,B.Y. and
            Pratt,J.Y.
  TITLE     Schizophrenia related genes
  JOURNAL   Patent: WO 0175440-A 15 11-OCT-2001;
            WELFIDE CORPORATION (JP)
FEATURES             Location/Qualifiers
     source          1..2697
                     /organism="Rattus sp."
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10118"
ORIGIN

Alignment Scores:
Pred. No.:              1.28e-279     Length:         2697
Score:                  4545.50       Matches:        852
Percent Similarity:     96.78%        Conservative:   19
Best Local Similarity:  94.67%        Mismatches:     26
Query Match:            94.96%        Indels:         3
DB:                     6             Gaps:           3

US-09-970-944-2 (1-899) x AX268596 (1-2697)

Qy          1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeu 20

Db          1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGTGCTCCTGGGCATAGTCCTCGCCGCCTGGCTT 60

Qy         21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40

Db         61 CGTGGTTCGGGTGCCCAGCAGAGTGCCACGGTGGCCAATCCAGTGCCCGGTGCCAACCCC 120

Qy         41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro 60

Db        121 GACCTGCTGCCCCACTTCCTGGTAGAGCCTGAGGACGTGTACATTGTCAAGAACAAGCCG 180

Qy         61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80

Db        181 GTGTTGTTGGTGTGCAAGGCTGTGCCTGCCACCCAGATCTTCTTCAAGTGCAATGGGGAA 240

Qy         81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100

Db        241 TGGGTCCGCCAGGTCGATCACGTAATTGAACGCAGCACCGACAGCAGCAGCGGATTGCCA 300

Qy        101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120

Db        301 ACCATGGAGGTCCGTATCAACGTATCGAGGCAGCAGGTAGAGAAAGTGTTTGGGCTGGAG 360

Qy        121 GluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAla 140

Db        361 GAATACTGGTGCCAGTGTGTGGCATGGAGCTCCTCGGGTACCACCAAAAGTCAGAAGGCC 420

Qy        141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160

Db        421 TACATCCGGATTGCCTATTTGCGCAAGAACTTTGAGCAGGAGCCACTGGCCAAGGAAGTG 480

Qy        161 SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu 180

Db        481 TCACTGGAGCAAGGCATTGTACTACCTTGTCGCCCCCCAGAAGGAATCCCCCCAGCTGAG 540

Qy        181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200

Db        541 GTGGAGTGGCTTCGAAATGAGGACCTCGTGGACCCCTCCCTCGATCCCAATGTGTACATC 600

Qy        201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220

Db        601 ACGCGGGAGCACAGCCTAGTCGTGCGTCAGGCCCGCCTGGCCGACACGGCCAACTACACC 660

Qy        221 CysValAlaLysAsnIleValAlaArgArgArgSerAlaSerAlaAlaValIleValTyr 240

Db        661 TGTGTGGCCAAGAACATCGTAGCCCGTCGCCGAAGCACCTCTGCAGCGGTCATTGTTTAT 720

Qy        241 ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg 260

Db        721 GTGAACGGTGGGTGGTCGACGTGGACTGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGT 780

Qy        261 GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe 280

Db        781 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCACCTCTCAACGGGGGCGCCTTC 840

Qy        281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300

Db        841 TGTGAGGGGCAGAATGTC---CAGAAAACAGCCTGCGCCACTCTGTGCCCAGTGGATGGG 897

Qy        301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320

Db        898 AGCTGGAGTTCGTGGAGTAAGTGGTCAGCCTGTGGGCTTGACTGCACCCACTGGCGGAGC 957

Qy        321 ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeu 340

Db        958 CGCGAGTGCTCTGACCCAGCACCCCGCAATGGAGGTGAGGAGTGTCGGGGTGCTGACCTG 1017

Qy        341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360

Db       1018 GACACCCGCAACTGTACCAGTGACCTCTGCCTGCACACCGCTTCTTGCCCCGAGGACGTG 1077

Qy        361 AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuValLeu 380

Db       1078 GCTCTCTACATCGGCCTTGTCGCTGTGGCTGTGTGCCTCTTCTTGCTGTTGCTGGCCCTT 1137

Qy        381 IleLeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIle 400

Db       1138 GGACTCATTTACTGTCGCAAGAAGGAAGGGCTGGACTCCGATGTGGCCGACTCGTCCATC 1197

Qy        401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420

Db       1198 CTCACCTCGGGCTTCCAGCCTGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCACCTG 1257

Qy        421 LeuThrIleGlnProAspLeuSer---ThrThrThrThrTyrGlnGlySerLeuCysPro 439

Db       1258 CTCACCATCCAGCCAGACCTCAGCACCACCACTACCACCTACCAGGGCAGTCTATGTTCG 1317

Qy        440 ArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu 459

Db       1318 AGGCAGGATGGACCCAGCCCCAAGTTCCAGCTCTCTAATGGTCACCTGCTCAGCCCACTG 1377

Qy        460 GlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheVal 479

Db       1378 GGGAGTGGCCGCCATACGTTGCACCACAGCTCACCCACCTCTGAGGCTGAGGACTTCGTC 1437

Qy        480 SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr 499

Db       1438 TCCCGCCTCTCCACCCAAAACTACTTTCGTTCCCTGCCCCGCGGCACCAGCAACATGGCC 1497
```

```
Qy    500 TyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeu 519
Db   1498 TACGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACGGGGATCAGCCTC 1557

Qy    520 LeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHis 539
Db   1558 CTCATACCCCCGGATGCCATCCCCCGAGGAAAGATCTACGAGATCTACCTCACACTGCAC 1617

Qy    540 LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal 559
Db   1618 AAGCCAGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCAGTCGTT 1677

Qy    560 SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys 579
Db   1678 AGCTGTGGGCCCCCAGGAGTCCTGCTCACCCGGCCAGTCATCCTTGCAATGGACCACTGT 1737

Qy    580 GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySer 599
Db   1738 GGAGAGCCCAGCCCTGACAGCTGGAGTCTGCGCCTCAAAAAGCAGTCCTGCGAGGGCAGT 1797

Qy    600 TrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGln 619
Db   1798 TGGGAG---GATGTGCTGCACCTTGGTGAGGAGTCACCTTCCCACCTCTACTACTGCCAG 1854

Qy    620 LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly 639
Db   1855 CTGGAGGCCGGGGCCTGCTATGTCTTCACGGAGCAGCTGGGCCGCTTTGCCCTGGTAGGA 1914

Qy    640 GluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAla 659
Db   1915 GAGGCCCTCAGCGTGGCTGCCACCAAGCGCCTCAGGCTCCTTCTGTTTGCTCCCGTGGCC 1974

Qy    660 CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu 679
Db   1975 TGTACGTCCCTTGAGTACAACATCCGAGTGTACTGCCTACACGACACCCACGACGCTCTC 2034

Qy    680 LysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgVal 699
Db   2035 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTAGGTGGACAGCTGATCCAGGAGCCTCGCGTC 2094

Qy    700 LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719
Db   2095 CTGCACTTCAAAGACAGTTACCACAACCTACGTCTCTCCATCCACGACGTGCCCAGCTCC 2154

Qy    720 LeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn 739
Db   2155 CTGTGGAAGAGCAAGCTACTTGTCAGCTACCAGGAGATCCCTTTTTACCACATCTGGAAC 2214

Qy    740 GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759
Db   2215 GGCACCCAGCAGTATCTGCACTGCACCTTCACCCTGGAGCGCATCAACGCCAGCACCAGC 2274

Qy    760 AspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIle 779
Db   2275 GACCTGGCCTGCAAGGTGTGGGTGTGGCAGGTGGAGGGAGATGGGCAGAGCTTCAACATC 2334

Qy    780 AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAla 799
Db   2335 AACTTCAACATCACTAAGGACACAAGGTTTGCTGAATTGTTGGCTCTGGAGAGTGAAGGG 2394

Qy    800 GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
Db   2395 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAAAAG 2454

Qy    820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839
Db   2455 ATCATCGCCAGTCTGGACCCACCCTGCAGCCGGGGCGCCGACTGGAGAACTCTAGCCCAG 2514

Qy    840 LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859
Db   2515 AAACTTCACCTGGACAGCCATCTTAGCTTCTTTGCCTCCAAGCCCAGCCCTACAGCCATG 2574

Qy    860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879
Db   2575 ATCCTCAACCTATGGGAGGCACGGCACTTCCCCAACGGCAACCTCGGCCAGCTGGCAGCA 2634

Qy    880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
Db   2635 GCTGTGGCCGGACTGGGCCAACCAGATGCTGGCCTCTTCACGGTGTCGGAGGCCGAGTGT 2694
```

```
RESULT 6
RNU87305
LOCUS       RNU87305                2697 bp    mRNA    linear   ROD 15-MAY-1997
DEFINITION  Rattus norvegicus transmembrane receptor Unc5H1 mRNA, complete cds.
ACCESSION   U87305
VERSION     U87305.1  GI:2055391
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 2697)
  AUTHORS   Leonardo,E.D., Hinck,L., Masu,M., Keino-Masu,K., Ackerman,S.L. and
            Tessier-Lavigne,M.
  TITLE     Vertebrate homologues of C. elegans UNC-5 are candidate netrin
            receptors
  JOURNAL   Nature 386 (6627), 833-838 (1997)
  MEDLINE   97271897
   PUBMED   9126742
REFERENCE   2  (bases 1 to 2697)
  AUTHORS   Leonardo,E.D., Hinck,L., Masu,M., Keino-Masu,K. and
            Tessier-Lavigne,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-JAN-1997) Anatomy, UCSF, 513 Parnassus, San
            Francisco, CA 94143-0452, USA
FEATURES             Location/Qualifiers
     source          1..2697
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10116"
                     /tissue_type="brain and ventral spinal cord"
                     /dev_stage="18 day embyo and 13 day embryo"
     CDS             1..2697
                     /codon_start=1
                     /product="transmembrane receptor UNC5H1"
                     /protein_id="AAB57678.1"
                     /db_xref="GI:2055392"
                     /translation="MAVRPGLWPVLLGIVLAAWLRGSGAQQSATVANPVPGANPDLLP
                     HFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDSSSGLPTM
                     EVRINVSRQQVEKVFGLEEYWCQCVAWSSSGTTKSQKAYIRIAYLRKNFEQEPLAKEV
                     SLEQGIVLPCRPPEGIPPAEVEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTAN
                     YTCVAKNIVARRRSTSAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLN
                     GGAFCEGQNVQKTACATLCPVDGSWSSWSKWSACGLDCTHWRSRECSDPAPRNGGEEC
                     RGADLDTRNCTSDLCLHTASCPEDVALYIGLVAVAVCLFLLLLALGLIYCRKKEGLDS
                     DVADSSILTSGFQPVSIKPSKADNPHLLTIQPDLSTTTTTYQGSLCSRQDGPSPKFQL
                     SNGHLLSPLGSGRHTLHHSSPTSEAEDFVSRLSTQNYFRSLPRGTSNMAYGTFNFLGG
                     RLMIPNTGISLLIPPDAIPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPVVSCGPPGV
                     LLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGSWEDVLHLGEESPSHLYYCQLEAGAC
                     YVFTEQLGRFALVGEALSVAATKRLRLLLFAPVACTSLEYNIRVYCLHDTHDALKEVV
                     QLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWKSKLLVSYQEIPFYHIWNGT
                     QQYLHCTFTLERINASTSDLACKVWVWQVEGDGQSFNINFNITKDTRFAELLALESEG
                     GVPALVGPSAFKIPFLIRQKIIASLDPPCSRGADWRTLAQKLHLDSHLSFFASKPSPT
                     AMILNLWEARHFPNGNLGQLAAAVAGLGQPDAGLFTVSEAEC"
ORIGIN

Alignment Scores:
Pred. No.:              1.28e-279       Length:         2697
Score:                  4545.50         Matches:        852
Percent Similarity:     96.78%          Conservative:   19
Best Local Similarity:  94.67%          Mismatches:     26
Query Match:            94.96%          Indels:         3
DB:                     10              Gaps:           3

US-09-970-944-2 (1-899) x RNU87305 (1-2697)
```

```
Qy      1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeu 20
Db      1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGTGCTCCTGGGCATAGTCCTCGCCGCCTGGCTT 60

Qy     21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40
Db     61 CGTGGTTCGGGTGCCCAGCAGAGTGCCACGGTGGCCAATCCAGTGCCCGGTGCCAACCCC 120

Qy     41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro 60
Db    121 GACCTGCTGCCCCACTTCCTGGTAGAGCCTGAGGACGTGTACATTGTCAAGAACAAGCCG 180

Qy     61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80
Db    181 GTGTTGTTGGTGTGCAAGGCTGTGCCTGCCACCCAGATCTTCTTCAAGTGCAATGGGGAA 240

Qy     81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100
Db    241 TGGGTCCGCCAGGTCGATCACGTAATTGAACGCAGCACCGACAGCAGCAGCGGATTGCCA 300

Qy    101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120
Db    301 ACCATGGAGGTCCGTATCAACGTATCGAGGCAGCAGGTAGAGAAAGTGTTTGGGCTGGAG 360

Qy    121 GluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAla 140
Db    361 GAATACTGGTGCCAGTGTGTGGCATGGAGCTCCTCGGGTACCACCAAAAGTCAGAAGGCC 420

Qy    141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160
Db    421 TACATCCGGATTGCCTATTTGCGCAAGAACTTTGAGCAGGAGCCACTGGCCAAGGAAGTG 480

Qy    161 SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu 180
Db    481 TCACTGGAGCAAGGCATTGTACTACCTTGTCGCCCCCCAGAAGGAATCCCCCCAGCTGAG 540

Qy    181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200
Db    541 GTGGAGTGGCTTCGAAATGAGGACCTCGTGGACCCCTCCCTCGATCCCAATGTGTACATC 600

Qy    201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220
Db    601 ACGCGGGAGCACAGCCTAGTCGTGCGTCAGGCCCGCCTGGCCGACACGGCCAACTACACC 660

Qy    221 CysValAlaLysAsnIleValAlaArgArgArgSerAlaSerAlaAlaValIleValTyr 240
Db    661 TGTGTGGCCAAGAACATCGTAGCCCGTCGCCGAAGCACCTCTGCAGCGGTCATTGTTTAT 720

Qy    241 ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg 260
Db    721 GTGAACGGTGGGTGGTCGACGTGGACTGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGT 780

Qy    261 GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe 280
Db    781 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCACCTCTCAACGGGGGCGCCTTC 840

Qy    281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300
Db    841 TGTGAGGGGCAGAATGTC---CAGAAAACAGCCTGCGCCACTCTGTGCCCAGTGGATGGG 897

Qy    301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320
Db    898 AGCTGGAGTTCGTGGAGTAAGTGGTCAGCCTGTGGGCTTGACTGCACCCACTGGCGGAGC 957

Qy    321 ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeu 340
Db    958 CGCGAGTGCTCTGACCCAGCACCCCGCAATGGAGGTGAGGAGTGTCGGGGTGCTGACCTG 1017

Qy    341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360
Db   1018 GACACCCGCAACTGTACCAGTGACCTCTGCCTGCACACCGCTTCTTGCCCCGAGGACGTG 1077

Qy    361 AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuLeuValLeu 380
Db   1078 GCTCTCTACATCGGCCTTGTCGCTGTGGCTGTGTGCCTCTTCTTGCTGTTGCTGGCCCTT 1137

Qy    381 IleLeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIle 400
Db   1138 GGACTCATTTACTGTCGCAAGAAGGAAGGGCTGGACTCCGATGTGGCCGACTCGTCCATC 1197

Qy    401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420
Db   1198 CTCACCTCGGGCTTCCAGCCTGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCACCTG 1257

Qy    421 LeuThrIleGlnProAspLeuSer---ThrThrThrThrTyrGlnGlySerLeuCysPro 439
Db   1258 CTCACCATCCAGCCAGACCTCAGCACCACCACTACCACCTACCAGGGCAGTCTATGTTCG 1317

Qy    440 ArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu 459
Db   1318 AGGCAGGATGGACCCAGCCCCAAGTTCCAGCTCTCTAATGGTCACCTGCTCAGCCCACTG 1377

Qy    460 GlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheVal 479
Db   1378 GGGAGTGGCCGCCATACGTTGCACCACAGCTCACCCACCTCTGAGGCTGAGGACTTCGTC 1437

Qy    480 SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr 499
Db   1438 TCCCGCCTCTCCACCCAAAACTACTTTCGTTCCCTGCCCCGCGGCACCAGCAACATGGCC 1497

Qy    500 TyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeu 519
Db   1498 TACGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACGGGGATCAGCCTC 1557

Qy    520 LeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHis 539
Db   1558 CTCATACCCCCGGATGCCATCCCCCGAGGAAAGATCTACGAGATCTACCTCACACTGCAC 1617

Qy    540 LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal 559
Db   1618 AAGCCAGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCAGTCGTT 1677

Qy    560 SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys 579
Db   1678 AGCTGTGGGCCCCCAGGAGTCCTGCTCACCCGGCCAGTCATCCTTGCAATGGACCACTGT 1737

Qy    580 GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySer 599
Db   1738 GGAGAGCCCAGCCCTGACAGCTGGAGTCTGCGCCTCAAAAAGCAGTCCTGCGAGGGCAGT 1797

Qy    600 TrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGln 619
Db   1798 TGGGAG---GATGTGCTGCACCTTGGTGAGGAGTCACCTTCCCACCTCTACTACTGCCAG 1854

Qy    620 LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly 639
Db   1855 CTGGAGGCCGGGGCCTGCTATGTCTTCACGGAGCAGCTGGGCCGCTTTGCCCTGGTAGGA 1914

Qy    640 GluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAla 659
Db   1915 GAGGCCCTCAGCGTGGCTGCCACCAAGCGCCTCAGGCTCCTTCTGTTTGCTCCCGTGGCC 1974

Qy    660 CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu 679
Db   1975 TGTACGTCCCTTGAGTACAACATCCGAGTGTACTGCCTACACGACACCCACGACGCTCTC 2034

Qy    680 LysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgVal 699
Db   2035 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTAGGTGGACAGCTGATCCAGGAGCCTCGCGTC 2094

Qy    700 LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719
Db   2095 CTGCACTTCAAAGACAGTTACCACAACCTACGTCTCTCCATCCACGACGTGCCCAGCTCC 2154

Qy    720 LeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn 739
```

```
Db       2155 CTGTGGAAGAGCAAGCTACTTGTCAGCTACCAGGAGATCCCTTTTTACCACATCTGGAAC 2214

Qy        740 GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759
Db       2215 GGCACCCAGCAGTATCTGCACTGCACCTTCACCCTGGAGCGCATCAACGCCAGCACCAGC 2274

Qy        760 AspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIle 779
Db       2275 GACCTGGCCTGCAAGGTGTGGGTGTGGCAGGTGGAGGGAGATGGGCAGAGCTTCAACATC 2334

Qy        780 AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAla 799
Db       2335 AACTTCAACATCACTAAGGACACAAGGTTTGCTGAATTGTTGGCTCTGGAGAGTGAAGGG 2394

Qy        800 GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
Db       2395 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAAAAG 2454

Qy        820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839
Db       2455 ATCATCGCCAGTCTGGACCCACCCTGCAGCCGGGGCGCCGACTGGAGAACTCTAGCCCAG 2514

Qy        840 LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859
Db       2515 AAACTTCACCTGGACAGCCATCTTAGCTTCTTTGCCTCCAAGCCCAGCCCTACAGCCATG 2574

Qy        860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879
Db       2575 ATCCTCAACCTATGGGAGGCACGGCACTTCCCCAACGGCAACCTCGGCCAGCTGGCAGCA 2634

Qy        880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
Db       2635 GCTGTGGCCGGACTGGGCCAACCAGATGCTGGCCTCTTCACGGTGTCGGAGGCCGAGTGT 2694
```

```
RESULT 7
BD057524
LOCUS       BD057524                3014 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Netrin receptors.
ACCESSION   BD057524
VERSION     BD057524.1  GI:22603130
KEYWORDS    JP 2001505062-A/1.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 3014)
  AUTHORS   Lavigne,M.T., Leonardo,D.E., Hinck,L., Masu,M. and Masu,K.K.
  TITLE     Netrin receptors
  JOURNAL   Patent: JP 2001505062-A 1 17-APR-2001;
            THE REGENTS OF THE UNIV OF CALIFORNIA
COMMENT     PN   JP 2001505062-A/1
            PD   17-APR-2001
            PF   19-FEB-1998 JP 1998536840
            PR   19-FEB-1997 US     08/808982
            PI   MARC TESSIER LAVIGNE,DAVID E LEONARDO,LINDSAY HINCK,MASAYUKI
            PI   MASU,
            PI   KAZUKO KEINO MASU
            PC   C07K1/00,C07K14/00,C07K17/00,C07H21/02,C07H21/04,G01N33/53 CC
             Strandedness: Double;
            CC   Topology: Linear;
            FH   Key           Location/Qualifiers.
FEATURES             Location/Qualifiers
     source          1..3014
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
ORIGIN

Alignment Scores:
Pred. No.:              1.48e-279      Length:          3014
Score:                  4545.50        Matches:         852
Percent Similarity:     96.78%         Conservative:    19
Best Local Similarity:  94.67%         Mismatches:      26
Query Match:            94.96%         Indels:          3
DB:                     6              Gaps:            3

US-09-970-944-2 (1-899) x BD057524 (1-3014)

Qy          1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeu 20
Db          1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGTGCTCCTGGGCATAGTCCTCGCCGCCTGGCTT 60

Qy         21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40
Db         61 CGTGGTTCGGGTGCCCAGCAGAGTGCCACGGTGGCCAATCCAGTGCCCGGTGCCAACCCC 120

Qy         41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro 60
Db        121 GACCTGCTGCCCCACTTCCTGGTAGAGCCTGAGGACGTGTACATTGTCAAGAACAAGCCG 180

Qy         61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80
Db        181 GTGTTGTTGGTGTGCAAGGCTGTGCCTGCCACCCAGATCTTCTTCAAGTGCAATGGGGAA 240

Qy         81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100
Db        241 TGGGTCCGCCAGGTCGATCACGTAATTGAACGCAGCACCGACAGCAGCAGCGGATTGCCA 300

Qy        101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120
Db        301 ACCATGGAGGTCCGTATCAACGTATCGAGGCAGCAGGTAGAGAAAGTGTTTGGGCTGGAG 360

Qy        121 GluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAla 140
Db        361 GAATACTGGTGCCAGTGTGTGGCATGGAGCTCCTCGGGTACCACCAAAAGTCAGAAGGCC 420

Qy        141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160
Db        421 TACATCCGGATTGCCTATTTGCGCAAGAACTTTGAGCAGGAGCCACTGGCCAAGGAAGTG 480

Qy        161 SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu 180
Db        481 TCACTGGAGCAAGGCATTGTACTACCTTGTCGCCCCCCAGAAGGAATCCCCCCAGCTGAG 540

Qy        181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200
Db        541 GTGGAGTGGCTTCGAAATGAGGACCTCGTGGACCCCTCCCTCGATCCCAATGTGTACATC 600

Qy        201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220
Db        601 ACGCGGGAGCACAGCCTAGTCGTGCGTCAGGCCCGCCTGGCCGACACGGCCAACTACACC 660

Qy        221 CysValAlaLysAsnIleValAlaArgArgArgSerAlaSerAlaAlaValIleValTyr 240
Db        661 TGTGTGGCCAAGAACATCGTAGCCCGTCGCCGAAGCACCTCTGCAGCGGTCATTGTTTAT 720

Qy        241 ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg 260
Db        721 GTGAACGGTGGGTGGTCGACGTGGACTGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGT 780

Qy        261 GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe 280
Db        781 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCACCTCTCAACGGGGGCGCCTTC 840

Qy        281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300
Db        841 TGTGAGGGGCAGAATGTC---CAGAAAACAGCCTGCGCCACTCTGTGCCCAGTGGATGGG 897

Qy        301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320
Db        898 AGCTGGAGTTCGTGGAGTAAGTGGTCAGCCTGTGGGCTTGACTGCACCCACTGGCGGAGC 957

Qy        321 ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeu 340
Db        958 CGCGAGTGCTCTGACCCAGCACCCCGCAATGGAGGTGAGGAGTGTCGGGGTGCTGACCTG 1017

Qy        341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360
```

```
                ||||||||||||||||||||||||||||||:::|||:::||||||    ||||||||||||
Db      1018 GACACCCGCAACTGTACCAGTGACCTCTGCCTGCACACCGCTTCTTGCCCCGAGGACGTG 1077

Qy       361 AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuLeuValLeu 380
             ||||||||:::||||||:::|||||||||||||||||    |||||||||||    |||
Db      1078 GCTCTCTACATCGGCCTTGTCGCTGTGGCTGTGTGCCTCTTCTTGCTGTTGCTGGCCCTT 1137

Qy       381 IleLeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIle 400
                |||:::||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1138 GGACTCATTTACTGTCGCAAGAAGGAAGGGCTGGACTCCGATGTGGCCGACTCGTCCATC 1197

Qy       401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1198 CTCACCTCGGGCTTCCAGCCTGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCACCTG 1257

Qy       421 LeuThrIleGlnProAspLeuSer---ThrThrThrThrTyrGlnGlySerLeuCysPro 439
             ||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
Db      1258 CTCACCATCCAGCCAGACCTCAGCACCACCACTACCACCTACCAGGGCAGTCTATGTTCG 1317

Qy       440 ArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu 459
             |||||||||||||||||||||||||||||||:::||||||||||||||||||||||||||
Db      1318 AGGCAGGATGGACCCAGCCCCAAGTTCCAGCTCTCTAATGGTCACCTGCTCAGCCCACTG 1377

Qy       460 GlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheVal 479
             |||    |||||||||||||||||||||||||||||||||||||||||||:::||||||
Db      1378 GGGAGTGGCCGCCATACGTTGCACCACAGCTCACCCACCTCTGAGGCTGAGGACTTCGTC 1437

Qy       480 SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr 499
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1438 TCCCGCCTCTCCACCCAAAACTACTTTCGTTCCCTGCCCCGCGGCACCAGCAACATGGCC 1497

Qy       500 TyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeu 519
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1498 TACGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACGGGGATCAGCCTC 1557

Qy       520 LeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHis 539
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1558 CTCATACCCCCGGATGCCATCCCCCGAGGAAAGATCTACGAGATCTACCTCACACTGCAC 1617

Qy       540 LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal 559
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db      1618 AAGCCAGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCAGTCGTT 1677

Qy       560 SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys 579
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1678 AGCTGTGGGCCCCCAGGAGTCCTGCTCACCCGGCCAGTCATCCTTGCAATGGACCACTGT 1737

Qy       580 GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySer 599
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1738 GGAGAGCCCAGCCCTGACAGCTGGAGTCTGCGCCTCAAAAAGCAGTCCTGCGAGGGCAGT 1797

Qy       600 TrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGln 619
             ||||||   ||||||||||||||||||||||||||:::||||||||||||||||||||||
Db      1798 TGGGAG---GATGTGCTGCACCTTGGTGAGGAGTCACCTTCCCACCTCTACTACTGCCAG 1854

Qy       620 LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly 639
             |||||||||    |||||||||||||||||||||||||||||||||||||||||||||||
Db      1855 CTGGAGGCCGGGGCCTGCTATGTCTTCACGGAGCAGCTGGGCCGCTTTGCCCTGGTAGGA 1914

Qy       640 GluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAla 659
             ||||||||||||||||||||      |||||||:::||||||||||||||||||||||||
Db      1915 GAGGCCCTCAGCGTGGCTGCCACCAAGCGCCTCAGGCTCCTTCTGTTTGCTCCCGTGGCC 1974

Qy       660 CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu 679
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1975 TGTACGTCCCTTGAGTACAACATCCGAGTGTACTGCCTACACGACACCCACGACGCTCTC 2034

Qy       680 LysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgVal 699
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2035 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTAGGTGGACAGCTGATCCAGGAGCCTCGCGTC 2094

Qy       700 LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2095 CTGCACTTCAAAGACAGTTACCACAACCTACGTCTCTCCATCCACGACGTGCCCAGCTCC 2154

Qy       720 LeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn 739
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2155 CTGTGGAAGAGCAAGCTACTTGTCAGCTACCAGGAGATCCCTTTTTACCACATCTGGAAC 2214

Qy       740 GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759
             ||||||||:::||||||||||||||||||||||||||||||||||::::::    |||||||||
Db      2215 GGCACCCAGCAGTATCTGCACTGCACCTTCACCCTGGAGCGCATCAACGCCAGCACCAGC 2274

Qy       760 AspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIle 779
             ||||||||||||||:::||||||||||||||||||||||||||||||||||||||:::|||
Db      2275 GACCTGGCCTGCAAGGTGTGGGTGTGGCAGGTGGAGGGAGATGGGCAGAGCTTCAACATC 2334

Qy       780 AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAla 799
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2335 AACTTCAACATCACTAAGGACACAAGGTTTGCTGAATTGTTGGCTCTGGAGAGTGAAGGG 2394

Qy       800 GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2395 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAAAAG 2454

Qy       820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839
             ||||||:::||||||||||||||||||||||      ||||||||||||||||||||||||||||
Db      2455 ATCATCGCCAGTCTGGACCCACCCTGCAGCCGGGGCGCCGACTGGAGAACTCTAGCCCAG 2514

Qy       840 LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2515 AAACTTCACCTGGACAGCCATCTTAGCTTCTTTGCCTCCAAGCCCAGCCCTACAGCCATG 2574

Qy       860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879
             ||||||||||||||||||||||||||||||||||||||||||||||    ||||||||||||
Db      2575 ATCCTCAACCTATGGGAGGCACGGCACTTCCCCAACGGCAACCTCGGCCAGCTGGCAGCA 2634

Qy       880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2635 GCTGTGGCCGGACTGGGCCAACCAGATGCTGGCCTCTTCACGGTGTCGGAGGCCGAGTGT 2694
```

RESULT 8
AX367094

```
LOCUS       AX367094                3580 bp    DNA     linear   PAT 16-FEB-2002
DEFINITION  Sequence 13 from Patent WO0198354.
ACCESSION   AX367094
VERSION     AX367094.1  GI:18855296
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Griffin,J.A., Kallick,D.A., Tribouley,C.M., Yue,H., Nguyen,D.B.,
            Tang,Y.T., Lal,P., Policky,J.L., Azimzai,Y., Lu,D.A., Graul,R.,
            Yao,M.G., Burford,N., Hafalia,A.J., Baughn,M.R., Bandman,O.,
            Patterson,C., Yang,J., Xu,Y., Warren,B.A., Ding,L. and
            Sanjanwala,M.S.
  TITLE     Receptors
  JOURNAL   Patent: WO 0198354-A 13 27-DEC-2001;
            Incyte Genomics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1..3580
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="Incyte ID No: 6052371CB1"
ORIGIN

Alignment Scores:
Pred. No.:              6.64e-269     Length:          3580
Score:                  4379.50       Matches:         838
Percent Similarity:     93.22%        Conservative:    1
Best Local Similarity:  93.11%        Mismatches:      2
Query Match:            91.49%        Indels:          59
DB:                     6             Gaps:            3
```

```
US-09-970-944-2 (1-899) x AX367094 (1-3580)

Qy          1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeu 20
Db          4 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 63

Qy         21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40
Db         64 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 123

Qy         41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro 60
Db        124 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 183

Qy         61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80
Db        184 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 243

Qy         81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100
Db        244 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 303

Qy        101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120
Db        304 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 363

Qy        121 GluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAla 140
Db        364 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 423

Qy        141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160
Db        424 TACATCCGCATAGCCTATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 483

Qy        161 SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu 180
Db        484 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 543

Qy        181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200
Db        544 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 603

Qy        201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220
Db        604 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 663

Qy        221 CysValAlaLysAsnIleValAlaArgArgArgSerAlaSerAlaAlaValIleValTyr 240
Db        664 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 723

Qy        241 ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg 260
Db        723 ------------------------------------------------------------ 723

Qy        261 GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe 280
Db        723 ------------------------------------------------------------ 723

Qy        281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300
Db        724 ---------------------------------------------------GTGGACGGC 732

Qy        301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320
Db        733 AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC 792

Qy        321 ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeu 340
Db        793 CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG 852

Qy        341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360
Db        853 GACACCCGCAACTGTACCAGTGACCTCTGTGTACACACTGCTTCTGGCCCTGAGGACGTG 912

Qy        361 AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuLeuValLeu 380
Db        913 GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC 972

Qy        381 IleLeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIle 400
Db        973 ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1032

Qy        401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420
Db       1033 CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1092

Qy        421 LeuThrIleGlnProAspLeuSer---ThrThrThrThrTyrGlnGlySerLeuCysPro 439
Db       1093 CTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCC 1152

Qy        440 ArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu 459
Db       1153 CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG 1212

Qy        460 GlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheVal 479
Db       1213 GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC 1272

Qy        480 SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr 499
Db       1273 TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC 1332

Qy        500 TyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeu 519
Db       1333 TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGAATCAGCCTC 1392

Qy        520 LeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHis 539
Db       1393 CTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC 1452

Qy        540 LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal 559
Db       1453 AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT 1512

Qy        560 SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys 579
Db       1513 AGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT 1572

Qy        580 GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySer 599
Db       1573 GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC 1632

Qy        600 TrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGln 619
Db       1633 TGGGAG---GATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG 1689

Qy        620 LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly 639
Db       1690 CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA 1749

Qy        640 GluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAla 659
Db       1750 GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC 1809

Qy        660 CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu 679
Db       1810 TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC 1869

Qy        680 LysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgVal 699
Db       1870 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTC 1929

Qy        700 LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719
Db       1930 CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCC 1989
```

```
Qy         720 LeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn 739
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1990 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2049

Qy         740 GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2050 GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2109

Qy         760 AspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIle 779
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2110 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2169

Qy         780 AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAla 799
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2170 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2229

Qy         800 GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2230 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2289

Qy         820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2290 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2349

Qy         840 LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2350 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG 2409

Qy         860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2410 ATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2469

Qy         880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2470 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2529

RESULT 9
BC058084
LOCUS       BC058084                3844 bp    mRNA    linear   ROD 21-OCT-2003
DEFINITION  Mus musculus unc-5 homolog A (C. elegans), mRNA (cDNA clone
            MGC:66671 IMAGE:6813463), complete cds.
ACCESSION   BC058084
VERSION     BC058084.1  GI:34784158
KEYWORDS    MGC.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 3844)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  MEDLINE   22388257
   PUBMED   12477932
REFERENCE   2  (bases 1 to 3844)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-SEP-2003) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Jim Lin, University of Iowa
            cDNA Library Preparation: M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Genome Sequence Centre,
            BC Cancer Agency, Vancouver, BC, Canada
            info@bcgsc.bc.ca
            Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
            Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
            Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
            Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
            Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
            Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
            Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
            George Yang, Scott Zuyderduyn, Marco Marra.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 126 Row: b Column: 11
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 23346570.
FEATURES             Location/Qualifiers
     source          1..3844
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="MGC:66671 IMAGE:6813463"
                     /tissue_type="Brain, mouse, 13.5, 14.5, 16.5, 17.5 dpc"
                     /clone_lib="NIH_BMAP_FW0"
                     /lab_host="DH10B"
                     /note="Vector: pYX-ASC"
     gene            1..3844
                     /gene="Unc5a"
                     /db_xref="LocusID:107448"
                     /db_xref="MGI:894682"
     CDS             252..2780
                     /codon_start=1
                     /product="Unc5a protein"
                     /protein_id="AAH58084.1"
                     /db_xref="GI:34784159"
                     /db_xref="LocusID:107448"
                     /translation="MAVRPGLWPALLGIVLTAWLRGSGAQQSATVANPVPGANPDLLP
                     HFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTM
                     EVRINVSRQQVEKVFGLEEYWCQCVAWSSSGTTKSQKAYIRIAYLRKNFEQEPLAKEV
                     SLEQGIVLPCRPPEGIPPAEVEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTPN
                     YTCVAKNIVARRRSASAAVIVYVDGSWSPWSKWSACGLDCTHWRSRECSDPAPRNGGE
                     ECRGADLDTRNCTSDLCLHTSSGPEDVALYIGLVAVAVCLILLLLVLVLIYCRKKEGL
                     DSDVADSSILTSGFQPVSIKPSKADNPHLLTIQPDLSTTTTTYQGSLCPRQDGPSPKF
                     QLSNGHLLSPLGSGRHTLHHSSPTSEAEDFVSRLSTQNYFRSLPRGTSNMAYGTFNFL
                     GGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPP
                     GVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGSWEDVLHLGEESPSHLYYCQLEAG
                     ACYVFTEQLGRFALVGEALSVAATKRLRLLLFAPVACTSLEYNIRVYCLHDTHDALKE
                     VVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWKSKLLVSYQEIPFYHIWN
                     GTQQYLHCTFTLERVNASTSDLACKVWVWQVEGDGQSFNINFNITKDTRFAEMLALES
                     EGGVPALVGPSAFKIPFLIRQKIITSLDPPCSRGADWRTLAQKLHLDSHLSFFASKPS
                     PTAMILNLWEARHFPNGNLGQLAAAVAGLGQPDAGLFTVSEAEC"
     misc_feature    714..926
                     /note="IG; Region: Immunoglobulin"
                     /db_xref="CDD:smart00409"
     misc_feature    984..1130
                     /note="TSP1; Region: Thrombospondin type 1 repeats"
                     /db_xref="CDD:smart00209"
     misc_feature    1566..1877
```

```
                  /note="ZU5; Region: Domain present in ZO-1 and Unc5-like
                  netrin receptors"
                  /db_xref="CDD:smart00218"
     misc_feature 2496..2684
                  /note="DEATH; Region: DEATH domain, found in proteins
                  involved in cell death (apoptosis). Alpha-helical domain
                  present in a variety of proteins with apoptotic functions.
                  Some (but not all) of these domains form homotypic and
                  heterotypic dimers"
                  /db_xref="CDD:smart00005"
ORIGIN

Alignment Scores:
Pred. No.:              1.93e-262     Length:         3844
Score:                  4278.50       Matches:        809
Percent Similarity:     92.00%        Conservative:   19
Best Local Similarity:  89.89%        Mismatches:     13
Query Match:            89.38%        Indels:         59
DB:                     10            Gaps:           3

US-09-970-944-2 (1-899) x BC058084 (1-3844)

Qy      1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeu 20
Db    252 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCACTGCCTGGCTT 311

Qy     21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40
Db    312 CGTGGTTCGGGTGCCCAGCAGAGTGCCACAGTGGCCAACCCAGTGCCTGGTGCCAACCCG 371

Qy     41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro 60
Db    372 GACCTGCTGCCCCACTTCCTGGTAGAGCCGGAGGACGTGTACATTGTCAAGAACAAGCCC 431

Qy     61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80
Db    432 GTGCTGCTGGTGTGCAAGGCTGTGCCCGCCACCCAGATCTTCTTCAAGTGCAACGGGGAA 491

Qy     81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100
Db    492 TGGGTTCGCCAGGTCGATCACGTCATTGAACGCAGCACTGACGGCAGCAGCGGATTGCCA 551

Qy    101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120
Db    552 ACCATGGAGGTCCGGATCAACGTATCAAGGCAGCAGGTCGAGAAAGTGTTTGGGCTGGAG 611

Qy    121 GluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAla 140
Db    612 GAGTACTGGTGCCAGTGTGTGGCATGGAGCTCCTCAGGAACCACCAAAAGCCAGAAGGCC 671

Qy    141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160
Db    672 TACATCCGGATTGCCTATTTGCGCAAGAACTTTGAGCAGGAGCCGCTGGCCAAGGAAGTG 731

Qy    161 SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu 180
Db    732 TCACTGGAGCAAGGCATTGTGCTACCTTGTCGCCCCCCGGAAGGAATCCCCCCAGCTGAG 791

Qy    181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200
Db    792 GTGGAGTGGCTCCGAAATGAGGACCTCGTGGACCCCTCCCTCGACCCCAATGTGTACATC 851

Qy    201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220
Db    852 ACACGGGAGCACAGCCTAGTCGTGCGGCAGGCCCGCCTGGCCGACACTCCCAACTACACC 911

Qy    221 CysValAlaLysAsnIleValAlaArgArgArgSerAlaSerAlaAlaValIleValTyr 240
Db    912 TGCGTGGCCAAGAACATCGTGGCCCGTCGCCGAAGCGCCTCTGCGGCCGTCATTGTTTAT 971

Qy    241 ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg 260
Db    971 ------------------------------------------------------------ 971

Qy    261 GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe 280
Db    971 ------------------------------------------------------------ 971

Qy    281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300
Db    972 ---------------------------------------------------GTGGATGGG 980

Qy    301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320
Db    981 AGCTGGAGCCCATGGAGTAAGTGGTCAGCCTGCGGGCTTGACTGCACCCACTGGCGGAGC 1040

Qy    321 ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeu 340
Db   1041 CGGGAGTGCTCCGACCCAGCGCCCCGCAACGGAGGTGAGGAGTGCCGGGGTGCTGACCTG 1100

Qy    341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360
Db   1101 GACACCCGCAACTGTACCAGTGACCTCTGCCTGCACACCTCTTCCGGCCCCGAGGACGTG 1160

Qy    361 AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuLeuValLeu 380
Db   1161 GCTCTCTACATCGGCCTCGTCGCCGTGGCCGTGTGCCTCATCTTGCTGCTGCTGGTCCTC 1220

Qy    381 IleLeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIle 400
Db   1221 GTCCTCATCTACTGCCGCAAGAAGGAAGGACTGGACTCAGACGTGGCTGACTCATCCATC 1280

Qy    401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420
Db   1281 CTTACCTCAGGCTTCCAGCCTGTCAGCATCAAGCCCAGCAAAGCAGACAATCCCCATCTG 1340

Qy    421 LeuThrIleGlnProAspLeuSer---ThrThrThrThrTyrGlnGlySerLeuCysPro 439
Db   1341 CTCACCATCCAACCGGACCTCAGCACCACCACGACCACCTACCAGGGCAGCCTGTGTCCC 1400

Qy    440 ArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu 459
Db   1401 CGGCAGGATGGACCCAGCCCCAAGTTCCAGCTCTCTAATGGTCACCTGCTCAGCCCACTG 1460

Qy    460 GlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheVal 479
Db   1461 GGCAGTGGCCGCCATACGCTGCACCACAGCTCCCCCACCTCTGAGGCTGAGGACTTCGTC 1520

Qy    480 SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr 499
Db   1521 TCCCGCCTCTCCACCCAAAACTACTTTCGTTCTCTGCCCCGCGGTACCAGCAACATGGCC 1580

Qy    500 TyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeu 519
Db   1581 TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAACACAGGAATCAGCCTC 1640

Qy    520 LeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHis 539
Db   1641 CTCATACCCCCGGACGCCATCCCCCGAGGAAAGATCTACGAGATCTACCTCACTCTGCAC 1700

Qy    540 LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal 559
Db   1701 AAGCCAGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCTATCGTT 1760

Qy    560 SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys 579
Db   1761 AGCTGTGGGCCCCCAGGAGTCCTGCTCACCCGGCCAGTCATCCTTGCCATGGACCACTGC 1820

Qy    580 GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySer 599
Db   1821 GGGGAGCCCAGTCCCGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCCTGTGAGGGCAGC 1880

Qy    600 TrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGln 619
Db   1881 TGGGAG---GACGTGCTGCACCTTGGTGAGGAGTCGCCCTCTCATCTCTACTACTGCCAG 1937
```

```
Qy        620 LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly 639
              ||||||||    |||||||||||||||||||||||||||||||||||||||||||||||||
Db       1938 CTGGAGGCCGGGGCCTGCTATGTCTTTACCGAGCAGCTAGGCCGCTTTGCCCTGGTGGGA 1997

Qy        640 GluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAla 659
              ||||||||||||||||||||     ||||||||:::||||||||||||||||||||||||
Db       1998 GAGGCCCTCAGCGTGGCTGCCACCAAGCGCCTCAGGCTCCTTCTGTTTGCCCCTGTGGCC 2057

Qy        660 CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu 679
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2058 TGTACGTCCCTCGAGTACAACATCCGAGTGTACTGCCTGCACGACACCCACGATGCTCTC 2117

Qy        680 LysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgVal 699
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2118 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGTGGACAGCTGATCCAGGAGCCCCGTGTC 2177

Qy        700 LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2178 CTGCACTTCAAAGACAGTTACCACAACCTACGTCTGTCCATCCACGACGTGCCCAGCTCC 2237

Qy        720 LeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn 739
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2238 CTGTGGAAGAGCAAGCTCCTTGTCAGCTACCAGGAGATCCCTTTTTACCACATCTGGAAT 2297

Qy        740 GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759
              ||||||||:::||||||||||||||||||||||||||||||||:::     |||||||||
Db       2298 GGCACTCAGCAGTATCTGCACTGCACCTTCACCCTGGAGCGCGTCAATGCCAGCACCAGC 2357

Qy        760 AspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIle 779
              ||||||||||||:::|||||||||||||||||||||||||||||||||||||:::|||
Db       2358 GACCTGGCCTGCAAGGTGTGGGTGTGGCAGGTGGAGGGAGATGGACAGAGCTTCAACATC 2417

Qy        780 AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAla 799
              |||||||||||||||||||||||||||||||||:::||||||||||||||||||
Db       2418 AACTTTAACATCACTAAGGACACGAGGTTTGCTGAAATGCTGGCTCTGGAGAGTGAAGGG 2477

Qy        800 GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2478 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAAAAG 2537

Qy        820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839
              ||||||:::||||||||||||||||||||||     ||||||||||||||||||||||||
Db       2538 ATCATTACCAGCCTGGACCCACCCTGCAGCCGGGGCGCCGACTGGCGAACTCTAGCCCAG 2597

Qy        840 LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2598 AAACTTCACCTGGACAGCCATCTTAGCTTCTTTGCCTCCAAGCCCAGCCCTACAGCCATG 2657

Qy        860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879
              |||||||||||||||||||||||||||||||||||||||||||     ||||||||||||
Db       2658 ATCCTCAACCTATGGGAGGCGCGGCACTTCCCCAACGGCAACCTCGGCCAGCTGGCCGCA 2717

Qy        880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2718 GCTGTGGCCGGACTGGGCCAGCCAGATGCTGGCCTCTTCACCGTGTCAGAGGCCGAGTGC 2777

RESULT 10
AK122575
LOCUS       AK122575                4294 bp    mRNA    linear   ROD 15-MAR-2003
DEFINITION  Mus musculus mRNA for mKIAA1976 protein.
ACCESSION   AK122575
VERSION     AK122575.1  GI:28972881
KEYWORDS    FLI_CDNA.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Okazaki,N., Kikuno,R., Ohara,R., Inamoto,S., Aizawa,H., Yuasa,S.,
            Nakajima,D., Nagase,T., Ohara,O. and Koga,H.
  TITLE     Prediction of the coding sequences of mouse homologues of KIAA
            gene: II. The complete nucleotide sequences of 400 mouse
            KIAA-homologous cDNAs identified by screening of terminal sequences
            of cDNA clones randomly sampled from size-fractionated libraries
  JOURNAL   DNA Res. 10, 35-48 (2003)
REFERENCE   2  (bases 1 to 4294)
  AUTHORS   Okazaki,N., Kikuno,R., Nagase,T., Ohara,O. and Koga,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-FEB-2003) Hisashi Koga, Kazusa DNA Research
            Institute, Laboratory for Genome Informatics; 2-6-7
            Kazusa-kamatari, Kisarazu, Chiba 292-0818, Japan
            (E-mail:mouse@kazusa.or.jp, Tel:81-438-52-3919, Fax:81-438-52-3918)
COMMENT     The CREATE program supported by Japan science and technology
            corporation; cDNA full insert sequencing: Kazusa DNA Research
            Institute; cDNA library construction, clone selection and 5'- &
            3'-end one pass sequencing.
FEATURES             Location/Qualifiers
     source          1..4294
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="mbh03993"
                     /tissue_type="brain"
                     /dev_stage="adult"
                     /note="vector:modified pBC SK+"
     gene            1..4294
                     /gene="mKIAA1976"
     CDS             <2263..3249
                     /gene="mKIAA1976"
                     /note="CDS is predicted by in silico analysis. Start codon
                     is not identified."
                     /codon_start=1
                     /evidence=not_experimental
                     /product="mKIAA1976 protein"
                     /protein_id="BAC65857.1"
                     /db_xref="GI:28972882"
                     /translation="SLYIPAILKTTLSLAASSGPEDVALYIGLVAVAVCLILLLLVLV
                     LIYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHLLTIQPDLSTTTTYQGSLC
                     PRQDGPSPKFQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWKSKLLVSYQE
                     IPFYHIWNGTQQYLHCTFTLERVNASTSDLACKVWVWQVEGDGQSFNINFNITKDTRF
                     AEMLALESEGGVPALVGPSAFKIPFLIRQKIITSLDPPCSRGADWRTLAQKLHLDSHL
                     SFFASKPSPTAMILNLWEARHFPNGNLGQLAAAVAGLGQPDAGLFTVSEAEC"
ORIGIN

Alignment Scores:
Pred. No.:              1.21e-182    Length:          4294
Score:                  3024.50      Matches:         632
Percent Similarity:     49.09%       Conservative:    15
Best Local Similarity:  47.95%       Mismatches:      15
Query Match:            63.18%       Indels:          657
DB:                     10           Gaps:            4

US-09-970-944-2 (1-899) x AK122575 (1-4294)

Qy          2 AlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeuArg 21
              ||||||||||||||||||||||||||||||||||||||||||||||     |||||||||||
Db          2 GCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCACTGCCTGGCTTCGT 61

Qy         22 GlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnProAsp 41
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 GGTTCGGGTGCCCAGCAGAGTGCCACAGTGGCCAACCCAGTGCCTGGTGCCAACCCGGAC 121

Qy         42 LeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysProVal 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 CTGCTGCCCCACTTCCTGGTAGAGCCGGAGGACGTGTACATTGTCAAGAACAAGCCCGTG 181

Qy         62 LeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGluTrp 81
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 CTGCTGGTGTGCAAGGCTGTGCCCGCCACCCAGATCTTCTTCAAGTGCAACGGGGAATGG 241

Qy         82 ValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluProThr 101
              ||||||||||||||||||||||||||||||||||||||||||||||||      ||||||
Db        242 GTTCGCCAGGTCGATCACGTCATTGAACGCAGCACTGACGGCAGCAGCGGATTGCCAACC 301

Qy        102 MetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGluGlu 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        302 ATGGAGGTCCGGATCAACGTATCAAGGCAGCAGGTCGAGAAAGTGTTTGGGCTGGAGGAG 361

Qy        122 TyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAlaTyr 141
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 TACTGGTGCCAGTGTGTGGCATGGAGCTCCTCAGGAACCACCAAAAGCCAGAAGGCCTAC 421

Qy        142 IleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluValSer 161
              |||||||||||      |||||||||||||||||||||||||||||||||||||||||||
Db        422 ATCCGGATTGCCTATTTGCGCAAGAACTTTGAGCAGGAGCCGCTGGCCAAGGAAGTGTCA 481

Qy        162 LeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGluVal 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 CTGGAGCAAGGCATTGTGCTACCTTGTCGCCCCCCGGAAGGAATCCCCCCAGCTGAGGTG 541

Qy        182 GluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIleThr 201
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 GAGTGGCTCCGAAATGAGGACCTCGTGGACCCCTCCCTCGACCCCAATGTGTACATCACA 601

Qy        202 ArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThrCys 221
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 CGGGAGCACAGCCTAGTCGTGCGGCAGGCCCGCCTGGCCGACACTGCCAACTACACCTGC 661

Qy        222 ValAlaLysAsnIleValAlaArgArgArgSerAlaSerAlaAlaValIleValTyrVal 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 GTGGCCAAGAACATCGTGGCCCGTCGCCGAAGCGCCTCTGCGGCCGTCATTGTTTATGTG 721

Qy        242 AsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArgGly 261
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        722 AACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGTGGC 781

Qy        262 TrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPheCys 281
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        782 TGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCACCTCTCAACGGGGGCGCCTTCTGT 841

Qy        282 GluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGlySer 301
              |||||||||||||      :::|||       :::    |||       |||||||||||
Db        842 GAGGGGCAGAATGTC---CAGAAAACAGCCTGCGCCACTCTGTGCCCAGTGGATGGGAGC 898

Qy        302 TrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSerArg 321
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        899 TGGAGCCCATGGAGTAAGTGGTCAGCCTGCGGGCTTGACTGCACCCACTGGCGGAGCCGG 958

Qy        322 GluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeuAsp 341
              ||||||||||||||||||||||||||||||||||||:::|||    ||||||||
Db        959 GAGTGCTCCGACCCAGCGCCCCGCAACGGAGGTGAGGAGTGCCGGGGTGCTGACCTGGAC 1018

Qy        342 ThrArgAsnCysThrSerAspLeuCysValHisSerAla--------------------- 354
              ||||||||||||||||||||||||||:::||||||
Db       1019 ACCCGCAACTGTACCAGTGACCTCTGCCTGCACAGTGA-GTCTTCTCTGCCCCAGGGCTC 1077

Qy        354 ------------------------------------------------------------ 354

Db       1078 ACCTCCGTCTGATAGCCCTATGGGCTGTGGTCAGGGCAGCCACTCTCCTGCCTGTGTCTC 1137

Qy        354 ------------------------------------------------------------ 354

Db       1138 TCCAAAGCTTGGGAGCCCAGGGGTCCCTTTCATCTCTAGCACACTCTTCCCTGTTACACA 1197

Qy        354 ------------------------------------------------------------ 354

Db       1198 GACACACCACCTTCCCCCTGAGACACACACAGATGTGAATCTTGACCCCTGTACATCCAA 1257

Qy        354 ------------------------------------------------------------ 354

Db       1258 CTGCTCCCCCGCGGCTTTGAAATTATCCCGTCCCCCAAGCTCACACATAAGTCCTGTGGG 1317

Qy        354 ------------------------------------------------------------ 354

Db       1318 GGTGCTGGGCTTGTGAGGCCTGTTAATTACAAGGCCATAACTAACAATTAAAGATCTCTG 1377

Qy        354 ------------------------------------------------------------ 354

Db       1378 TTGGATTTTTTTTTTCCTTTAGACTACACCCTCCCCTGTTCTTTCAATTAAAAACCTAA 1437

Qy        354 ------------------------------------------------------------ 354

Db       1438 TTATCTGTGTCAATTCCTTTCTGATGGACATTTCTTCCATTATGCAAAACTCTTAAAAAT 1497

Qy        354 ------------------------------------------------------------ 354

Db       1498 CGATAGTGCTTCCTCCAAGCAGTTGGTCCCCCACCCACCCCAGTTTCAGGGCTGGTGGTA 1557

Qy        354 ------------------------------------------------------------ 354

Db       1558 TACACAAGGGCTCACACTGTCCTGGAAACACCCGAGAAACTGACCACCCATGTATCTGCC 1617

Qy        354 ------------------------------------------------------------ 354

Db       1618 TATTTTCTATTTCTATAATGAACTAGCAAAAGAAATTTATTAGTTTTAGAGGCTCAAATT 1677

Qy        354 ------------------------------------------------------------ 354

Db       1678 CCAAATAGCGTGATGCTGGCTCTGGCAAATGTGTCTTCCACCCCAAGCCACATTGCTTGG 1737

Qy        354 ------------------------------------------------------------ 354

Db       1738 TGGCAGATAGTACCACATCAGTACTGTAGACAAGAGAGAGATCACGTGTCAGGAAGAGAG 1797

Qy        354 ------------------------------------------------------------ 354

Db       1798 AAAGACAGACACAGAGACAGGTAGACACGGAAAGATGGACAGCCCCATGGTCTAAGAGCA 1857

Qy        354 ------------------------------------------------------------ 354

Db       1858 GCACTCATAATGATGTAATGCCTCCCCACTGGGGCCCACCTCCTAAAAGGCCCCACTGCT 1917

Qy        354 ------------------------------------------------------------ 354

Db       1918 GCTCATATTGCCACAATGGGGACCACACCTGTCAATATGAATCCTTAGGCAGGCATGTGC 1977

Qy        354 ------------------------------------------------------------ 354

Db       1978 AAACTATAGCACCATACTTAGGGTCTCTCAGGAGTTCACCTTGCCCCAGACACCCTGACC 2037

Qy        354 ------------------------------------------------------------ 354

Db       2038 CCCAGCCTAGGAGCCAAGGATATACCACCACTCACCCCGGTGAAGTAGCTCTTCCCCTGC 2097

Qy        354 ------------------------------------------------------------ 354

Db       2098 TCCACTCTGCCTGTGAGGGTGTGACTTGCTCAATCCCAAGCAAGGTTGTGGTGGGTATAG 2157

Qy        354 ------------------------------------------------------------ 354

Db       2158 AACTAGGATGTACCCACCAATACCACGCTCCCTGCCACGTGCCCACTGCCCAGGGGGTGC 2217

Qy        354 ------------------------------------------------------------ 354

Db       2218 TGAGATCCTATTCTCCTGGGGATAAACCATGGGCCTCTCCTCTGAAGCCTCTACATCCCT 2277

Qy        355 ------------------------------------SerGlyProGluAspValAlaLeu 362
                                                  ||||||||||||||||||||||||
Db       2278 GCCATTCTCAAGACCACACTTTCTCTCGCAGCCTCTTCCGGCCCCGAGGACGTGGCTCTC 2337

Qy        363 TyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuLeuValLeuIleLeu 382
              |||:::||||||:::||||||||||||||||||:::||||||||||||||||||:::|||
Db       2338 TACATCGGCCTCGTCGCCGTGGCCGTGTGCCTCATCTTGCTGCTGCTGGTCCTCGTCCTC 2397

Qy        383 ValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIleLeuThr 402
              :::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2398 ATCTACTGCCGCAAGAAGGAAGGACTGGACTCAGATGTGGCTGACTCATCCATCCTTACC 2457

Qy        403 SerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeuLeuThr 422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2458 TCAGGCTTCCAGCCTGTCAGCATCAAGCCCAGCAAAGCAGACAATCCCCATCTGCTCACC 2517
```

```
Qy      423 IleGlnProAspLeuSer---ThrThrThrThrTyrGlnGlySerLeuCysProArgGln 441
            ||||||||||||||||||   |||||||||||||||||||||||||||||||||||||||
Db     2518 ATCCAACCGGACCTCAGCACCACCACGACCACCTACCAGGGCAGCCTGTGTCCCCGGCAG 2577

Qy      442 AspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeuGlyGly 461
            ||||||||||||||||||||
Db     2578 GATGGACCCAGCCCCAAGTTC--------------------------------------- 2598

Qy      462 GlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheValSerArg 481

Db     2598 ------------------------------------------------------------ 2598

Qy      482 LeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThrTyrGly 501

Db     2598 ------------------------------------------------------------ 2598

Qy      502 ThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeuLeuIle 521

Db     2598 ------------------------------------------------------------ 2598

Qy      522 ProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHisLysPro 541

Db     2598 ------------------------------------------------------------ 2598

Qy      542 GluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleValSerCys 561

Db     2598 ------------------------------------------------------------ 2598

Qy      562 GlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGlyGlu 581

Db     2598 ------------------------------------------------------------ 2598

Qy      582 ProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySerTrpGlu 601

Db     2598 ------------------------------------------------------------ 2598

Qy      602 GlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGlnLeuGlu 621

Db     2598 ------------------------------------------------------------ 2598

Qy      622 AlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGluAla 641

Db     2598 ------------------------------------------------------------ 2598

Qy      642 LeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAlaCysThr 661

Db     2598 ------------------------------------------------------------ 2598

Qy      662 SerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuLysGlu 681

Db     2598 ------------------------------------------------------------ 2598

Qy      682 ValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgValLeuHis 701
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2599 ------CAGCTGGAGAAGCAGCTGGGTGGACAGCTGATCCAGGAGCCCCGTGTCCTGCAC 2652

Qy      702 PheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSerLeuTrp 721
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2653 TTCAAAGACAGTTACCACAACCTACGTCTGTCCATCCACGACGTGCCCAGCTCCCTGTGG 2712

Qy      722 LysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsnGlyThr 741
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2713 AAGAGCAAGCTCCTTGTCAGCTACCAGGAGATCCCTTTTTACCACATCTGGAATGGCACT 2772

Qy      742 GlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSerAspLeu 761
            |||:::||||||||||||||||||||||||||||||||:::     ||||||||||||||
Db     2773 CAGCAGTATCTGCACTGCACCTTCACCCTGGAGCGCGTCAATGCCAGCACCAGCGACCTG 2832

Qy      762 AlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIleAsnPhe 781
            |||||||||:::||||||||||||||||||||||||||||||||||||:::|||||||||
Db     2833 GCCTGCAAGGTGTGGGTGTGGCAGGTGGAGGGAGATGGACAGAGCTTCAACATCAACTTT 2892

Qy      782 AsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAlaGlyVal 801
            ||||||||||||||||||||||||||||:::||||||||||||||||||   ||||||
Db     2893 AACATCACTAAGGACACGAGGTTTGCTGAAATGCTGGCTCTGGAGAGTGAAGGGGGGGTC 2952

Qy      802 ProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLysIleIle 821
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2953 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAAAAGATCATT 3012

Qy      822 SerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGlnLysLeu 841
            :::||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db     3013 ACCAGCCTGGACCCACCCTGCAGCCGGGGCGCTGACTGGCGAACTCTAGCCCAGAAACTT 3072

Qy      842 HisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMetIleLeu 861
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3073 CACCTGGACAGCCATCTTAGCTTCTTTGCCTCCAAGCCCAGCCCTACAGCCATGATCCTC 3132

Qy      862 AsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAlaAlaVal 881
            ||||||||||||||||||||||||||||||||||||   |||||||||||||||||||||
Db     3133 AACCTATGGGAGGCGCGGCACTTCCCCAACGGCAACCTCGGCCAGCTGGCCGCAGCTGTG 3192

Qy      882 AlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3193 GCCGGACTGGGCCAGCCAGATGCTGGCCTCTTCACCGTGTCAGAGGCCGAGTGC 3246
```

```
RESULT 11
BC009333
LOCUS       BC009333                2688 bp    mRNA    linear   PRI 19-DEC-2003
DEFINITION  Homo sapiens cDNA clone IMAGE:4126760, partial cds.
ACCESSION   BC009333
VERSION     BC009333.2  GI:40226527
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2688)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
REFERENCE   2  (bases 1 to 2688)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JUN-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     On Dec 19, 2003 this sequence version replaced gi:14424611.
            Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC),
            Gaithersburg, Maryland;
```

```
              Web site: http://www.nisc.nih.gov/
              Contact: nisc_mgc@nhgri.nih.gov
              Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
              Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
              Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
              Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
              Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D.,McCloskey,J.C.,
              McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
              Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
              Young,A., Zhang,L.-H. and Green,E.D.

              Clone distribution: MGC clone distribution information can be found
              through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
              Series: IRAL Plate: 26 Row: g Column: 22.
FEATURES             Location/Qualifiers
     source          1..2688
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4126760"
                     /tissue_type="Brain, neuroblastoma"
                     /clone_lib="NIH_MGC_19"
                     /lab_host="DH10B-R"
                     /note="Vector: pOTB7"
     CDS             <1..1627
                     /codon_start=2
                     /product="Unknown (protein for IMAGE:4126760)"
                     /protein_id="AAH09333.2"
                     /db_xref="GI:40226528"
                     /translation="DVALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILT
                     SGFQPVSIKPSKADNPHLLTIQPDLSTTTTTYQGSLCPRQDGPSPKFQLTNGHLLSPL
                     GGGRHTLHHSSPTSEAEEFVSRLSTQNYFRSLPRGTSNMTYGTFNFLGGRLMIPNTGI
                     SLLIPPDAIPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILA
                     MDHCGEPSPDSWSLRLKKQSCEGSWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGR
                     FALVGEALSVAAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQ
                     LIQEPRVLHFKDSYHNLRLSIHDVPSSLWKSKLLVSYQEIPFYHIWNGTQRYLHCTFT
                     LERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPS
                     AFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLDSHLSFFASKPSPTAMILNLWEA
                     RHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC"
     misc_feature    413..724
                     /note="ZU5; Region: Domain present in ZO-1 and Unc5-like
                     netrin receptors"
                     /db_xref="CDD:smart00218"
     misc_feature    1343..1594
                     /note="DEATH; Region: DEATH domain, found in proteins
                     involved in cell death (apoptosis)"
                     /db_xref="CDD:smart00005"
ORIGIN

Alignment Scores:
Pred. No.:              7.68e-170      Length:         2688
Score:                  2819.00        Matches:        540
Percent Similarity:     99.63%         Conservative:   0
Best Local Similarity:  99.63%         Mismatches:     0
Query Match:            58.89%         Indels:         2
DB:                     9              Gaps:           2

US-09-970-944-2 (1-899) x BC009333 (1-2688)

Qy        359 AspValAlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuLeu 378
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 GACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTT 61

Qy        379 ValLeuIleLeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSer 398
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 GTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCG 121

Qy        399 SerIleLeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnPro 418
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 TCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCC 181

Qy        419 HisLeuLeuThrIleGlnProAspLeuSer---ThrThrThrThrTyrGlnGlySerLeu 437
              ||||||||||||||||||||||||||||||   |||||||||||||||||||||||||||
Db        182 CATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTC 241

Qy        438 CysProArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSer 457
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 TGTCCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGC 301

Qy        458 ProLeuGlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGlu 477
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 CCCCTGGGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAG 361

Qy        478 PheValSerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsn 497
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 TTCGTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAAC 421

Qy        498 MetThrTyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIle 517
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 ATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGAATC 481

Qy        518 SerLeuLeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThr 537
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 AGCCTCCTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACG 541

Qy        538 LeuHisLysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerPro 557
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 CTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCC 601

Qy        558 IleValSerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAsp 577
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 ATCGTTAGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGAC 661

Qy        578 HisCysGlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGlu 597
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 CACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAG 721

Qy        598 GlySerTrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyr 617
              ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db        722 GGCAGCTGGGAG---GATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTAC 778

Qy        618 CysGlnLeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeu 637
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        779 TGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTG 838

Qy        638 ValGlyGluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaPro 657
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        839 GTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCG 898

Qy        658 ValAlaCysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAsp 677
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        899 GTGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGAT 958

Qy        678 AlaLeuLysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluPro 697
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        959 GCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCA 1018

Qy        698 ArgValLeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValPro 717
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1019 CGGGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCC 1078

Qy        718 SerSerLeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIle 737
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1079 AGCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATC 1138

Qy        738 TrpAsnGlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSer 757
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1139 TGGAATGGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGC 1198

Qy        758 ThrSerAspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPhe 777
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1199 ACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTC 1258

Qy        778 SerIleAsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSer 797
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1259 AGCATCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGT 1318
```

```
Qy        798 GluAlaGlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArg 817
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1319 GAAGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGG 1378

Qy        818 GlnLysIleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeu 837
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1379 CAGAAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTG 1438

Qy        838 AlaGlnLysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThr 857
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1439 GCCCAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACA 1498

Qy        858 AlaMetIleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeu 877
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1499 GCCATGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTG 1558

Qy        878 AlaAlaAlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAla 897
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1559 GCTGCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCT 1618

Qy        898 GluCys 899
              ||||||
Db       1619 GAGTGC 1624
```

```
RESULT 12
MMU72634
LOCUS       MMU72634                9299 bp    mRNA    linear   ROD 13-MAY-1997
DEFINITION  Mus musculus rostral cerebellar malformation protein (rcm) mRNA,
            complete cds.
ACCESSION   U72634
VERSION     U72634.1  GI:2088526
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 9299)
  AUTHORS   Ackerman,S.L., Kozak,L.P., Przyborski,S.A., Rund,L.A., Boyer,B.B.
            and Knowles,B.B.
  TITLE     The mouse rostral cerebellar malformation gene encodes an
            UNC-5-like protein
  JOURNAL   Nature 386 (6627), 838-842 (1997)
  MEDLINE   97271898
   PUBMED   9126743
REFERENCE   2  (bases 1 to 9299)
  AUTHORS   Ackerman,S.L., Kozak,L.P., Rund,L.A. and Knowles,B.B.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-SEP-1996) The Jackson Laboratory, 600 Main Street,
            Bar Harbor, ME 04609, USA
FEATURES             Location/Qualifiers
     source          1..9299
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57B6/SJL"
                     /db_xref="taxon:10090"
                     /chromosome="3"
                     /map="between Amy1 and Mpmv9"
     gene            1..9299
                     /gene="rcm"
     CDS             151..2946
                     /gene="rcm"
                     /function="transmembrane receptor"
                     /note="similar to C. elegans UNC-5"
                     /codon_start=1
                     /product="rostral cerebellar malformation protein"
                     /protein_id="AAB54103.1"
                     /db_xref="GI:2088527"
                     /translation="MRKGLRATAARCGLGLGYLLQMLVLPALALLSASGTGSAAQDDE
                     FFHELPETFPSDPPEPLPHFLIEPEEAYIVKNKPVNLYCKASPATQIYFKCNSEWVHQ
                     KDHVVDERVDETSGLIVREVSIEISRQQVEELFGPEDYWCQCVAWSSAGTTKSRKAYV
                     RIAYLRKTFEQEPLGKEVSLEQEVLLQCRPPEGIPVAEVEWLKNEDIIDPAEDRNFYI
                     TIDHNLIIKQARLSDTANYTCVAKNIVAKRKSTTATVIVYVNGGWSTWTEWSVCNSRC
                     GRGYQKRTRTCTNPAPLNGGAFCEGQSVQKIACTTLCPVDGRWTSWSKWSTCGTECTH
                     WRRRECTAPAPKNGGKDCDGLVLQSKNCTDGLCMQAAPDSDDVALYVGIVIAVTVCLA
                     ITVVVALFVYRKNHRDFESDIIDSSALNGGFQPVNIKAARQDLLAVPPDLTSAAAMYR
                     GPVYALHDVSDKIPMTNSPILDPLPNLKIKVYNSSGAVTPQDDLAEFSSKLSPQMTQS
                     LLENEALNLKNQSLARQTDPSCTAFGTFNSLGGHLIIPNSGVSLLIPAGAIPQGRVYE
                     MYVTVHRKENMRPPMEDSQTLLTPVVSCGPPGALLTRPVILTLHHCADPSTEDWKIQL
                     KNQAVQGQWEDVVVVGEENFTTPCYIQLDAEACHILTENLSTYALVGQSTTKAAAKRL
                     KLAIFGPLCCSSLEYSIRVYCLDDTQDALKEVLQLERQMGGQLLEEPKALHFKGSIHN
                     LRLSIHDIAHSLWKSKLLAKYQEIPFYHIWSGSQRNLHCTFTLERLSLNTVELVCKLC
                     VRQVEGEGQIFQLNCTVSEEPTGIDLPLLDPASTITTVTGPSAFSIPLPIRQKLCSSL
                     DAPQTRGHDWRMLAHKLNLDRYLNYFATKSSPTGVILDLWEAQNFPDGNLSMLAAVLE
                     EMGRHETVVSLAAEGQY"
ORIGIN
```

```
Alignment Scores:
Pred. No.:                 6.4e-162     Length:          9299
Score:                     2705.50      Matches:         515
Percent Similarity:        73.17%       Conservative:    153
Best Local Similarity:     56.41%       Mismatches:      214
Query Match:               56.52%       Indels:          31
DB:                        10           Gaps:            12
```

```
US-09-970-944-2 (1-899) x MMU72634 (1-9299)

Qy          9 ProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeuArgGlySerGlyAlaGlnGlnSer 28
              ||||||||          :::|||:::|||       ||||||   ||||||
Db        226 CCTGCCCTG------GCCCTGCTAAGCGCCAGTGGCACCGGCTCCGCCGCTCAAGATGAT 279

Qy         29 Ala---------ThrValAlaAsnProValProGlyAlaAsnProAspLeuLeuProHis 45
                         :::         |||         |||:::   |||||||||
Db        280 GAATTTTTTCACGAACTCCCAGAAACCTTTCCATCTGACCCACCTGAGCCATTGCCACAC 339

Qy         46 PheLeuValGluProGluAspValTyrIleValLysAsnLysProValLeuLeuValCys 65
              ||||||:::|||||||||:::   ||||||||||||||||||||||||   |||   |||
Db        340 TTCCTCATTGAGCCCGAGGAAGCTTACATTGTGAAGAACAAGCCTGTGAACCTGTATTGT 399

Qy         66 LysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGluTrpValArgGlnVal 85
              ||||||   |||||||||||||:::||||||||||||   |||||||||   |||
Db        400 AAAGCCAGCCCTGCCACCCAGATCTACTTCAAGTGCAACAGCGAGTGGGTTCATCAGAAG 459

Qy         86 AspHisValIleGluArgSerThrAspGlySerSerGlyGluProThrMetGluValArg 105
              |||||||||::::::         |||   :::||||||         ||||||
Db        460 GACCACGTAGTAGACGAGAGAGTAGATGAAACCTCTGGTCTAATTGTGAGAGAAGTGAGC 519

Qy        106 IleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGluGluTyrTrpCysGln 125
              |||   :::|||||||||||||||||||||::::::||||||   |||:::||||||||||
Db        520 ATTGAGATTTCACGCCAGCAGGTGGAGGAACTGTTTGGGCCTGAAGATTACTGGTGCCAG 579

Qy        126 CysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAlaTyrIleArgIleAla 145
              ||||||||||||||||||:::||||||||||||||||||:::|||||||||:::|||||||
Db        580 TGTGTGGCCTGGAGCTCAGCAGGCACTACGAAGAGTCGGAAGGCATACGTGCGCATTGCG 639

Qy        146 ArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluValSerLeuGluGlnGly 165
                 |||||||||      |||||||||||||||||||||||||||||||||
Db        640 TATCTGCGGAAGACATTCGAGCAGGAACCCTTGGGAAAGGAAGTGTCCTTGGAGCAGGAA 699

Qy        166 IleValLeuProCysArgProProGluGlyIleProProAlaGluValGluTrpLeuArg 185
              ::::::|||   |||||||||||||||||||||||   |||||||||||||||||:::
Db        700 GTCTTACTCCAGTGTCGGCCACCTGAAGGGATCCCAGTGGCTGAGGTGGAATGGCTAAAG 759

Qy        186 AsnGluAspLeuValAspProSerLeuAspProAsnValTyrIleThrArgGluHisSer 205
              |||||||||::::::||||||:::   |||   |||   |||||||||   :::|||:::
Db        760 AATGAAGACATAATTGATCCTGCTGAAGATCGGAACTTTTATATTACTATCGATCACAAC 819

Qy        206 LeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThrCysValAlaLysAsn 225
              |||:::::::::||||||||||||:::||||||||||||||||||||||||||||||||||
Db        820 CTGATCATCAAGCAAGCCCGACTCTCAGATACAGCAAATTATACCTGTGTTGCCAAAAAT 879

Qy        226 IleValAlaArgArgArgSerAlaSerAlaAlaValIleValTyrValAsnGlyGlyTrp 245
              |||||||||:::|||:::|||      :::|||   ||||||||||||||||||||||||
Db        880 ATTGTTGCCAAGAGAAAAAGCACCACAGCCACTGTCATCGTGTATGTTAATGGTGGCTGG 939

Qy        246 SerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArgGlyTrpGlnLysArg 265
```

```
                  ||||||||||||||||||||||||||::::::   |||||||||||:::|||||||||
Db        940 TCCACCTGGACAGAGTGGTCTGTGTGTAACAGCCGCTGTGGGCGAGGATATCAGAAACGC 999

Qy        266 SerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPheCysGluGlyGlnAsn 285
              :::|||:::||||||||||||||||||||||||||||||||||||||||||||||||:::
Db       1000 ACAAGAACCTGCACCAACCCAGCCCCACTCAATGGTGGGGCCTTCTGTGAGGGGCAGAGT 1059

Qy        286 ValHisAspArgThrValSerSerLeuLeuValSerValAspGlySerTrpSerProTrp 305
              |||         :::   |||      ||||||||    |||:::      |||
Db       1060 GTG---CAGAAAATAGCATGCACTACGTTATGTCCAGTGGATGGTAGGTGGACTTCATGG 1116

Qy        306 SerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSerArgGluCysSerAsp 325
              |||||||||||   ||||||     :::||||||||||||||||     |||||||||:::
Db       1117 AGCAAATGGTCAACCTGTGGGACTGAATGCACCCACTGGCGCAGGAGGGAGTGTACAGCA 1176

Qy        326 ProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeuAspThrArgAsnCys 345
              ||||||||:::|||||||||::::::|||   |||      |||   ::::::||||||
Db       1177 CCAGCCCCCAAGAACGGGGGTAAGGACTGTGATGGCCTGGTCCTCCAATCCAAGAACTGC 1236

Qy        346 ThrSerAspLeuCysValHisSerAlaSerGlyProGluAspValAlaLeuTyrValGly 365
              |||       ||||||:::   :::|||          :::||||||||||||||||||
Db       1237 ACTGATGGGCTGTGCATGCAGGCTGCTCCTGACTCAGATGATGTGGCTCTCTACGTGGGG 1296

Qy        366 Leu---IleAlaValAlaValCysLeuValLeuLeuLeuLeuValLeuIleLeuValTyr 384
              :::   |||||||||       :::     ::::::|||     :::     ||||||
Db       1297 ATTGTGATCGCTGTAACAGTCTGTCTGGCGATCACTGTTGTGGTGGCCCTGTTTGTGTAT 1356

Qy        385 CysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIleLeuThrSerGly 404
                 :::                 :::||||||:::   |||||||||   |||      |||
Db       1357 CGGAAGAACCACCGTGACTTTGAGTCTGACATCATTGACTCCTCAGCACTCAATGGCGGC 1416

Qy        405 PheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeuLeuThrIleGln 424
              ||||||||||||:::||||||   ::::::   |||            ||||||   :::
Db       1417 TTTCAGCCTGTGAACATCAAGGCTGCCAGACAAGAT---------CTCCTGGCTGTCCCC 1467

Qy        425 ProAspLeu---SerThrThrThrThrTyrGlnGlySerLeuCysProArgGlnAspGly 443
              |||||||||   |||              |||:::|||   :::            |||
Db       1468 CCTGACCTCACCTCAGCTGCAGCCATGTACAGGGGACCTGTCTATGCTCTGCATGAT--- 1524

Qy        444 ProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeuGlyGlyGlyArg 463
                 |||   |||      :::||||||      :::|||   ||||||           :::
Db       1525 GTCTCAGACAAAATCCCAATGACCAACTCTCCAATTCTGGACCCACTACCCAACTTGAAA 1584

Qy        464 HisThrLeuHisHisSerSer------------ProThrSerGluAlaGluGluPheVal 479
                 :::::::::||||||             |||      :::      ||||||
Db       1585 ATCAAAGTGTACAACAGCTCAGGTGCTGTCACTCCTCAGGATGACCTTGCCGAGTTCTCA 1644

Qy        480 SerArgLeuSer---------ThrGlnAsnTyrPhe------------------------ 488
              |||:::||||||         ||||||:::
Db       1645 TCCAAACTGTCACCCCAGATGACCCAGTCCTTGCTAGAGAATGAGGCCCTTAACCTGAAG 1704

Qy        489 ---ArgSerLeuProArgGlyThr------SerAsnMetThrTyrGlyThrPheAsnPhe 505
                 :::||||||   |||   |||         |||         :::||||||||||||
Db       1705 AACCAGAGCCTCGCAAGACAGACTGACCCATCCTGCACAGCATTTGGTACCTTCAACTCT 1764

Qy        506 LeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeuLeuIleProProAspAla 525
              |||||||||   |||:::|||||||||:::|||:::||||||||||||||      |||
Db       1765 CTTGGGGGTCACCTCATCATTCCTAATTCAGGAGTAAGCTTGCTGATTCCCGCTGGGGCC 1824

Qy        526 IleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHisLysProGluAspValArg 545
              ||||||:::|||::::::|||:::|||:::|||:::      |||::::::|||
Db       1825 ATTCCTCAGGGGAGAGTCTATGAAATGTATGTGACTGTACACAGGAAAGAAAATATGAGG 1884

Qy        546 LeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleValSerCysGlyProProGly 565
                 |||:::         ||||||||||||:::|||:::|||||||||||||||||||||
Db       1885 CCCCCCATGGAAGACTCTCAGACCCTACTTACCCCTGTGGTGAGCTGTGGGCCTCCTGGA 1944

Qy        566 ValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGlyGluProSerProAsp 585
                 ||||||||||||||||||||||   :::   ||||||   :::||||||   :::
Db       1945 GCTCTGCTGACCCGCCCTGTCATCCTCACTCTGCATCACTGTGCAGACCCCAGCACCGAG 2004

Qy        586 SerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySerTrpGluGlnAspValLeu 605
                 |||   ::::::||||||   |||:::   :::|||   ||||||   ||||||:::

Db       2005 GACTGGAAGATCCAGCTCAAAAACCAGGCAGTGCAGGGACAATGGGAG---GATGTTGTG 2061

Qy        606 HisLeuGlyGluAlaProSerHisLeuTyrTyrCysGlnLeuGluAlaSerAlaCys 625
                 :::|||||||||         :::         |||   ||||||:::|||   ||||||
Db       2062 GTGGTTGGGGAGGAGAACTTCACAACCCCCTGTTACATTCAGCTGGATGCAGAGGCTTGC 2121

Qy        626 TyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGluAlaLeuSerValAla 645
              ::::::   ||||||   |||         :::||||||||||||::::::   :::   |||
Db       2122 CATATCCTCACAGAGAACCTCAGTACCTATGCCCTGGTTGGGCAGTCCACCACCAAAGCA 2181

Qy        646 AlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAlaCysThrSerLeuGluTyr 665
              ||||||||||||||||||||   :::|||   |||:::   |||:::||||||||||||
Db       2182 GCTGCCAAGCGTCTTAAACTGGCCATCTTTGGGCCCCTCTGCTGCTCTTCCCTGGAGTAC 2241

Qy        666 AsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuLysGluValValGlnLeu 685
              :::||||||||||||||||||   ||||||   ||||||||||||||||||:::||||||
Db       2242 AGCATTAGAGTCTACTGCCTGGATGACACACAGGATGCCCTGAAGGAAGTTCTACAACTG 2301

Qy        686 GluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgValLeuHisPheLysAspSer 705
              |||:::|||:::||||||||||||::::::||||||:::   ||||||||||||   |||
Db       2302 GAGAGGCAAATGGGAGGACAGCTCCTAGAAGAACCCAAGGCTCTTCATTTAAAGGCAGC 2361

Qy        706 TyrHisAsnLeuArgLeuSerIleHisAspValProSerSerLeuTrpLysSerLysLeu 725
                 |||||||||||||||||||||||||:::         |||||||||||||||||||
Db       2362 ATCCACAACCTGCGCCTGTCTATTCATGACATCGCCCATTCCCTCTGGAAGAGCAAATTG 2421

Qy        726 LeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsnGlyThrGlnArgTyrLeu 745
              |||         |||||||||||||||||||||||||||:::|||:::||||||   |||
Db       2422 CTGGCTAAGTATCAGGAAATTCCATTTTACCACATCTGGAGTGGCTCTCAAAGAAACCTC 2481

Qy        746 HisCysThrPheThrLeuGluArgValSerProSerThrSerAspLeuAlaCysLysLeu 765
              ||||||||||||||||||||||||:::|||   :::|||   :::|||   |||||||||
Db       2482 CACTGCACCTTCACTCTGGAAAGACTCAGCCTAAACACAGTGGAACTGGTTTGCAAACTC 2541

Qy        766 TrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIleAsnPheAsnIleThrLys 785
                 |||   ||||||||||||:::||||||   |||   :::|||         ::::::::
Db       2542 TGTGTGCGGCAGGTTGAAGGAGAAGGGCAGATCTTCCAGCTCAACTGTACTGTGTCAGAG 2601

Qy        786 AspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAlaGlyValProAlaLeuVal 805
              :::         :::|||      |||:::      :::   :::      :::
Db       2602 GAACCTACTGGCATCGACTTACCTCTCCTGGACCCTGCTAGTACCATCACCACTGTCACC 2661

Qy        806 GlyProSerAlaPheLysIleProPheLeuIleArgGlnLysIleIleSerSerLeuAsp 825
              |||||||||||||||   ||||||      ||||||||||||:::   ||||||||||||
Db       2662 GGACCAAGTGCTTTCAGCATTCCTCTCCCTATCCGGCAGAAGCTATGCAGCAGCCTGGAT 2721

Qy        826 ProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGlnLysLeuHisLeuAspSer 845
              |||         ||||||   |||||||||   ||||||   ||||||:::||||||
Db       2722 GCCCCTCAAACAAGAGGCCATGACTGGAGGATGCTGGCCCATAAACTCAACCTGGACAGG 2781

Qy        846 HisLeuSerPhePheAlaSerLysProSerProThrAlaMetIleLeuAsnLeuTrpGlu 865
              :::|||::::::||||||:::|||   |||||||||   :::||||||:::|||||||||
Db       2782 TACTTGAATTACTTTGCCACCAAATCGAGCCCAACTGGCGTAATCCTGGATCTTTGGGAA 2841

Qy        866 AlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAlaValAlaGlyLeuGly 885
              |||:::::::|||||||:::||||||||||||   |||||||||   :::      :::|||
Db       2842 GCACAGAACTTCCCAGATGGAAACCTGAGCATGCTGGCAGCCGTCCTGGAAGAAATGGGA 2901

Qy        886 GlnProAspAlaGlyLeuPheThrValSerGluAlaGlu 898
              :::   :::      :::         :::|||   :::
Db       2902 AGACATGAGACAGTGGTGTCCTTGGCAGCAGAAGGACAG 2940
```

```
RESULT 13
BD057525
LOCUS       BD057525                1787 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Netrin receptors.
ACCESSION   BD057525
VERSION     BD057525.1  GI:22603131
KEYWORDS    JP 2001505062-A/2.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 1787)
```

```
AUTHORS    Lavigne,M.T., Leonardo,D.E., Hinck,L., Masu,M. and Masu,K.K.
TITLE      Netrin receptors
JOURNAL    Patent: JP 2001505062-A 2 17-APR-2001;
           THE REGENTS OF THE UNIV OF CALIFORNIA
COMMENT    PN   JP 2001505062-A/2
           PD   17-APR-2001
           PF   19-FEB-1998 JP 1998536840
           PR   19-FEB-1997 US    08/808982
           PI   MARC TESSIER LAVIGNE,DAVID E LEONARDO,LINDSAY HINCK,MASAYUKI
           PI   MASU,
           PI   KAZUKO KEINO MASU
           PC   C07K1/00,C07K14/00,C07K17/00,C07H21/02,C07H21/04,G01N33/53 CC
            Strandedness: Double;
           CC   Topology: Linear;
           FH   Key             Location/Qualifiers.
FEATURES             Location/Qualifiers
     source          1. .1787
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
ORIGIN

Alignment Scores:
Pred. No.:              4e-162        Length:         1787
Score:                  2694.00       Matches:        543
Percent Similarity:     97.14%        Conservative:   1
Best Local Similarity:  96.96%        Mismatches:     11
Query Match:            56.28%        Indels:         11
DB:                     6             Gaps:           2

US-09-970-944-2 (1-899) x BD057525 (1-1787)

Qy        344 AsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspValAlaLeuTyr 363
              ||||||||||||||||    |||||:::||||||||||||||||||||||||||||||
Db          3 AACTGTACCAGTGACCTCTG-GTACACACTGCTTCTGGCCCTGAGGACGTGGCCCTCTAT 61

Qy        364 ValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuLeuValLeuIleLeuVal 383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 GTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCATCCTCGTT 121

Qy        384 TyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIleLeuThrSer 403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 TATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCA 181

Qy        404 GlyPheGlnProValSerIle-LysProSerLysAlaAspAsnProHisLeuLeuThrIl 423
              ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        182 GGCTTCCAGCCCGTCAGCATCTAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCAT 241

Qy        423 eGlnProAspLeuSer---ThrThrThrThrTyrGlnGlySerLeuCysProArgGlnAs 442
              ||||||||||||||||   |||||||||||||||||||||||||||||||||||||||||
Db        242 CCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGA 301

Qy        442 pGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeuGlyGlyGl 462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 TGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGG 361

Qy        462 yArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheValSerArgLe 482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 CCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCT 421

Qy        482 uSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThrTyrGlyTh 502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 CTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGAC 481

Qy        502 rPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeuLeuIlePr 522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 CTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGAATCAGCCTCCTCATCCC 541

Qy        522 oProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHisLysProGl 542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 CCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGA 601

Qy        542 uAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleValSerCysGl 562
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 AGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGG 661

Qy        562 yProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGlyGluPr 582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 ACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCC 721

Qy        582 oSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySerTrpGluGl 602
              ||||||||||||||||||||     ||||||||||||||||||||||    ||||||
Db        722 CAGCCCTGACAGCTGGAGCCTGGCCCTCAAAAAGCAGTCGTGCGAGGGAGC-TGGGAG-- 778

Qy        602 nAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGlnLeuGluAl 622
               |||  |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        779 -GATGT-CTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGC 836

Qy        622 aSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGluAlaLe 642
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        837 CAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCT 896

Qy        642 uSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAlaCysThrSe 662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        897 CAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTC 956

Qy        662 rLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuLysGluVa 682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        957 CCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGT 1016

Qy        682 lValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgValLeuHisPh 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1017 GGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTT 1076

Qy        702 eLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSerLeuTrpLy 722
               |||||||||||||||||||    |||    ||||||||||||||||||||||||||||
Db       1077 -AAGGACAGTTACCACAACCTGCC-CTATC-ATCCACGATGTGCCCAGCTCCCTGTGGAA 1133

Qy        722 sSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsnGlyThrGl 742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1134 GAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCA 1193

Qy        742 nArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSerAspLeuAl 762
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1194 GCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGC 1253

Qy        762 aCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIleAsnPheAs 782
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1254 CTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAA 1313

Qy        782 nIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAlaGlyValPr 802
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1314 CATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCC 1373

Qy        802 oAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLysIleIleSe 822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1374 AGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTC 1433

Qy        822 rSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGlnLysLeuHi 842
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1434 CAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCA 1493

Qy        842 sLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMetIleLeuAs 862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1494 CCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAA 1553

Qy        862 nLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAlaAlaValAl 882
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1554 CCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGC 1613

Qy        882 aGlyLeuGlyGlnProAspAlaGlyLeu-PheThrVal-SerGluAlaGluCys 899
              ||||   ||||||   |||   |||    ||||||||| ||||||||||||||
Db       1614 TGGGACTGGCCAGCAGGACGGTGGCTTCTTTCACAGTGTTCGGAGGCTGAGTGC 1667

RESULT 14
AB118026
```

```
LOCUS       AB118026                9328 bp    mRNA    linear   ROD 19-DEC-2003
DEFINITION  Rattus norvegicus Unc5h3 mRNA, complete cds.
ACCESSION   AB118026
VERSION     AB118026.1  GI:40217509
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1
  AUTHORS   Kuramoto,T., Kuwamura,M. and Serikawa,T.
  TITLE     Rat neurological mutations cerebellar vermis defect and hobble are
            caused by mutations in the netrin-1 receptor gene Unc5h3
  JOURNAL   Mol. Brain Res. (2003) In press
REFERENCE   2  (bases 1 to 9328)
  AUTHORS   Kuramoto,T. and Serikawa,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-AUG-2003) Takashi Kuramoto, Institute of Laboratory
            Animals, Graduate School of Medicine, Kyoto University;
            Yoshidakonoe-cho, Sakyo-ku, Kyoto 606-8501, Japan
            (E-mail:tkuramot@anim.med.kyoto-u.ac.jp,
            URL:www.anim.med.kyoto-u.ac.jp, Tel:81-75-753-4494,
            Fax:81-75-753-4409)
FEATURES             Location/Qualifiers
     source          1..9328
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10116"
                     /chromosome="2"
     gene            1..9328
                     /gene="Unc5h3"
     CDS             151..2946
                     /gene="Unc5h3"
                     /codon_start=1
                     /protein_id="BAD05181.1"
                     /db_xref="GI:40217510"
                     /translation="MRKGLRATAARCGLGLGYLLQMLVLPALALLSASGTGSAAQDDD
                     FFHELPETFPSDPPEPLPHFLIEPEEAYIVKNKPVNLYCKASPATQIYFKCNSEWVHQ
                     KDHVVDERVDETSGLIVREVSIEISRQQVEELFGPEDYWCQCVAWSSAGTTKSRKAYV
                     RIAYLRKTFEQEPLGKEVSLEQEVLLQCRPPEGIPMAEVEWLKNEDIIDPVEDRNFYI
                     TIDHNLIIKQARLSDTANYTCVAKNIVAKRKSTTATVIVYVNGGWSTWAEWSVCNSRC
                     GRGYQKRTRTCTNPAPLNGGAFCEGQSVQKIACTTLCPVDGRWTSWSKWSTCGTECTH
                     WRRRECTAPAPKNGGKDCDGLVLQSKNCTDGLCMQAAPDSDDVALYVGIVIAVTVCLA
                     ITVVVALFVYRKNHRDFESNIIDSSALNGGFQPVNIKAARQDLLAVPPDLTSAAAMYR
                     GPVYALHDVSDKIPMTNSPILDPLPNLKIKVYNSSGAVTPQDDLAEFSSKLSPQMTQS
                     LLENEALNLKNQSLARQTDPSCTAFGTFNSLGGHLIIPNSGVSLLIPAGAIPQGRVYE
                     MYVTVHRKENMRPPMEDSQTLLTPVVSCGPPGALLTRPVILTLHHCADPNTEDWKIQL
                     KNQAVQGQWEDVVVVGEENFTTPCYIQLDAEACHILTENLSTYALVGQSTTKAAAKRL
                     KLAIFGPLCCSSLEYSIRVYCLDDTQDALKEVLQLERQMGGQLLEEPKALHFKGSIHN
                     LRLSIHDITHSLWKSKLLAKYQEIPFYHIWSGSQRNLHCTFTLERLSLNTVELVCKLC
                     VRQVEGEGQIFQLNCTVSEEPTGIDLPLLDPASTITTVTGPSAFSIPLPIRQKLCSSL
                     DAPQTRGHDWRMLAHKLNLDRYLNYFATKSSPTGVILDLWEAQNFPDGNLSMLAAVLE
                     EMGRHETVVSLAAEGQY"
     3'UTR           2947..9328
                     /gene="Unc5h3"
     repeat_region   8727..8772
                     /note="simple sequence repeats"
                     /rpt_unit="gt"
     polyA_signal    9309..9314
                     /gene="Unc5h3"
ORIGIN

Alignment Scores:
Pred. No.:              7.74e-161      Length:         9328
Score:                  2688.50        Matches:        512
Percent Similarity:     72.95%         Conservative:   154
Best Local Similarity:  56.08%         Mismatches:     216
Query Match:            56.16%         Indels:         31
DB:                     10             Gaps:           12

US-09-970-944-2 (1-899) x AB118026 (1-9328)
```

```
Qy      9 ProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeuArgGlySerGlyAlaGlnGlnSer 28
Db    226 CCTGCCCTG------GCCCTGCTCAGCGCCAGTGGCACCGGCTCGGCCGCTCAAGATGAT 279

Qy     29 ---------AlaThrValAlaAsnProValProGlyAlaAsnProAspLeuLeuProHis 45
Db    280 GACTTTTTCCATGAACTCCCAGAAACCTTTCCATCAGACCCACCTGAGCCATTACCACAC 339

Qy     46 PheLeuValGluProGluAspValTyrIleValLysAsnLysProValLeuLeuValCys 65
Db    340 TTTCTCATTGAGCCCGAGGAAGCTTACATTGTGAAGAACAAGCCTGTGAACCTGTATTGT 399

Qy     66 LysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGluTrpValArgGlnVal 85
Db    400 AAAGCCAGCCCTGCCACCCAGATCTACTTCAAGTGCAACAGCGAATGGGTTCATCAGAAG 459

Qy     86 AspHisValIleGluArgSerThrAspGlySerSerGlyGluProThrMetGluValArg 105
Db    460 GACCACGTAGTAGACGAGAGGGTAGATGAAACCTCCGGTCTCATTGTGAGAGAAGTGAGC 519

Qy    106 IleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGluGluTyrTrpCysGln 125
Db    520 ATTGAGATTTCACGCCAGCAGGTGGAGGAACTGTTCGGGCCTGAAGATTACTGGTGCCAG 579

Qy    126 CysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAlaTyrIleArgIleAla 145
Db    580 TGTGTGGCCTGGAGCTCAGCGGGCACTACGAAGAGTCGGAAGGCATACGTGCGCATTGCG 639

Qy    146 ArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluValSerLeuGluGlnGly 165
Db    640 TATCTGCGGAAGACATTTGAGCAGGAACCCTTGGGAAAAGAAGTGTCCTTGGAGCAGGAA 699

Qy    166 IleValLeuProCysArgProProGluGlyIleProProAlaGluValGluTrpLeuArg 185
Db    700 GTGTTACTCCAGTGCCGGCCACCTGAAGGGATCCCAATGGCTGAGGTGGAGTGGCTAAAG 759

Qy    186 AsnGluAspLeuValAspProSerLeuAspProAsnValTyrIleThrArgGluHisSer 205
Db    760 AATGAAGACATAATTGATCCTGTCGAAGATCGGAATTTTTATATCACTATCGATCACAAC 819

Qy    206 LeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThrCysValAlaLysAsn 225
Db    820 CTGATCATCAAGCAGGCACGACTCTCAGATACAGCGAATTATACCTGTGTTGCCAAAAAC 879

Qy    226 IleValAlaArgArgSerAlaSerAlaAlaValIleValTyrValAsnGlyGlyTrp 245
Db    880 ATTGTTGCCAAGAGGAAAAGCACCACAGCCACCGTCATCGTGTATGTTAATGGTGGCTGG 939

Qy    246 SerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArgGlyTrpGlnLysArg 265
Db    940 TCCACCTGGGCAGAATGGTCTGTGTGTAACAGCCGCTGTGGGCGAGGATATCAGAAACGC 999

Qy    266 SerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPheCysGluGlyGlnAsn 285
Db   1000 ACAAGAACCTGCACCAACCCAGCCCCCCTCAATGGTGGGGCCTTCTGCGAGGGGCAGAGT 1059

Qy    286 ValHisAspArgThrValSerSerLeuLeuValSerValAspGlySerTrpSerProTrp 305
Db   1060 GTA---CAGAAAATAGCATGCACCACGTTATGCCCAGTGGATGGCAGGTGGACTTCGTGG 1116

Qy    306 SerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSerArgGluCysSerAsp 325
Db   1117 AGCAAATGGTCAACTTGTGGGACTGAGTGCACCCACTGGCGGAGGAGGGAGTGTACAGCA 1176

Qy    326 ProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeuAspThrArgAsnCys 345
Db   1177 CCAGCCCCCAAGAACGGGGGTAAGGACTGTGATGGCCTGGTCCTCCAATCCAAGAACTGC 1236

Qy    346 ThrSerAspLeuCysValHisSerAlaSerGlyProGluAspValAlaLeuTyrValGly 365
Db   1237 ACCGATGGGCTGTGCATGCAGGCTGCTCCTGACTCAGATGACGTTGCTCTCTACGTTGGG 1296

Qy    366 Leu---IleAlaValAlaValCysLeuValLeuLeuLeuValLeuIleLeuValTyr 384
```

```
Db       1297 ATTGTGATCGCTGTAACCGTTTGTCTGGCAATCACTGTTGTGGTGGCCCTGTTCGTGTAT 1356
Qy        385 CysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIleLeuThrSerGly 404
Db       1357 CGCAAGAACCACCGTGACTTTGAGTCTAATATCATTGACTCCTCAGCACTCAATGGCGGC 1416
Qy        405 PheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeuLeuThrIleGln 424
Db       1417 TTTCAGCCTGTGAACATCAAGGCTGCCAGACAAGAT---------CTCCTGGCAGTACCC 1467
Qy        425 ProAspLeu---SerThrThrThrThrTyrGlnGlySerLeuCysProArgGlnAspGly 443
Db       1468 CCGGATCTCACCTCAGCTGCAGCCATGTACAGGGGACCCGTCTATGCTCTCCATGAT--- 1524
Qy        444 ProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeuGlyGlyGlyArg 463
Db       1525 GTCTCAGACAAAATCCCAATGACCAATTCTCCAATTCTAGACCCACTGCCCAATTTGAAA 1584
Qy        464 HisThrLeuHisHisSerSer------------ProThrSerGluAlaGluGluPheVal 479
Db       1585 ATCAAAGTGTACAACAGCTCAGGTGCTGTCACTCCTCAGGACGACCTTGCTGAGTTCTCA 1644
Qy        480 SerArgLeuSer---------ThrGlnAsnTyrPhe------------------------ 488
Db       1645 TCCAAGCTGTCACCCCAGATGACCCAATCCTTGCTAGAGAATGAGGCCCTGAACCTGAAG 1704
Qy        489 ---ArgSerLeuProArgGlyThr------SerAsnMetThrTyrGlyThrPheAsnPhe 505
Db       1705 AACCAGAGTCTCGCAAGACAGACTGACCCATCCTGCACAGCGTTTGGCACCTTCAACTCC 1764
Qy        506 LeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeuLeuIleProProAspAla 525
Db       1765 CTGGGCGGTCACCTCATCATCCCTAATTCAGGAGTAAGCTTGCTGATTCCCGCTGGGGCC 1824
Qy        526 IleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHisLysProGluAspValArg 545
Db       1825 ATTCCTCAGGGGAGAGTCTATGAAATGTATGTGACTGTACACAGGAAAGAAAATATGAGG 1884
Qy        546 LeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleValSerCysGlyProProGly 565
Db       1885 CCCCCCATGGAAGACTCTCAGACCCTGCTGACCCCTGTGGTGAGCTGTGGGCCTCCTGGA 1944
Qy        566 ValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGlyGluProSerProAsp 585
Db       1945 GCTCTGTTGACCCGTCCTGTCATCCTCACTCTGCATCACTGTGCAGACCCTAACACCGAG 2004
Qy        586 SerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySerTrpGluGlnAspValLeu 605
Db       2005 GACTGGAAGATCCAGCTCAAAAACCAGGCAGTGCAGGGACAGTGGGAG---GATGTTGTG 2061
Qy        606 HisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGlnLeuGluAlaSerAlaCys 625
Db       2062 GTTGTTGGAGAGGAAAACTTCACAACCCCCTGTTACATTCAGCTGGATGCAGAGGCTTGC 2121
Qy        626 TyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGluAlaLeuSerValAla 645
Db       2122 CATATTCTCACGGAGAACCTCAGTACCTATGCCCTGGTTGGGCAGTCCACCACCAAAGCA 2181
Qy        646 AlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAlaCysThrSerLeuGluTyr 665
Db       2182 GCTGCCAAGCGTCTGAAGCTGGCCATCTTTGGGCCCCTCTGCTGCTCTTCCCTGGAGTAT 2241
Qy        666 AsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuLysGluValValGlnLeu 685
Db       2242 AGCATCAGAGTCTACTGCCTGGATGACACACAGGATGCCCTGAAGGAAGTTCTACAACTG 2301
Qy        686 GluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgValLeuHisPheLysAspSer 705
Db       2302 GAAAGGCAAATGGGAGGACAGCTCCTAGAAGAACCCAAGGCTCTTCATTTTAAAGGCAGC 2361
Qy        706 TyrHisAsnLeuArgLeuSerIleHisAspValProSerSerLeuTrpLysSerLysLeu 725
Db       2362 ATCCACAACCTGCGCCTCTCAATTCATGACATCACCCATTCCCTCTGGAAGAGCAAATTG 2421
Qy        726 LeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsnGlyThrGlnArgTyrLeu 745
Db       2422 CTGGCTAAGTATCAGGAAATTCCATTTTACCACATCTGGAGTGGCTCTCAAAGAAACCTC 2481
Qy        746 HisCysThrPheThrLeuGluArgValSerProSerThrSerAspLeuAlaCysLysLeu 765
Db       2482 CACTGTACCTTCACTCTGGAAAGACTCAGCCTAAACACAGTAGAGCTGGTTTGCAAACTC 2541
Qy        766 TrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIleAsnPheAsnIleThrLys 785
Db       2542 TGTGTGCGGCAGGTTGAAGGAGAAGGGCAGATCTTCCAGCTCAACTGTACTGTGTCAGAG 2601
Qy        786 AspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAlaGlyValProAlaLeuVal 805
Db       2602 GAACCTACTGGTATCGATTTGCCTCTCCTGGACCCTGCTAGCACCATTACCACTGTCACT 2661
Qy        806 GlyProSerAlaPheLysIleProPheLeuIleArgGlnLysIleIleSerSerLeuAsp 825
Db       2662 GGACCAAGTGCTTTCAGCATCCCTCTTCCTATCCGGCAGAAGCTATGTAGCAGCCTGGAT 2721
Qy        826 ProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGlnLysLeuHisLeuAspSer 845
Db       2722 GCCCCCCAAACAAGAGGCCATGACTGGAGGATGCTGGCCCACAAACTTAACCTGGACAGG 2781
Qy        846 HisLeuSerPhePheAlaSerLysProSerProThrAlaMetIleLeuAsnLeuTrpGlu 865
Db       2782 TACTTGAATTACTTTGCCACCAAATCGAGCCCGACTGGCGTAATCCTGGATCTCTGGGAA 2841
Qy        866 AlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAlaAlaValAlaGlyLeuGly 885
Db       2842 GCACAGAACTTCCCCGATGGAAACCTGAGCATGCTGGCAGCCGTCCTGGAAGAAATGGGA 2901
Qy        886 GlnProAspAlaGlyLeuPheThrValSerGluAlaGlu 898
Db       2902 AGACACGAAACCGTAGTGTCCTTGGCAGCAGAAGGTCAG 2940
```

RESULT 15
AC123700/c

```
LOCUS       AC123700                200968 bp    DNA     linear   HTG 13-MAR-2003
DEFINITION  Mus musculus clone RP23-335A11, WORKING DRAFT SEQUENCE, 9 unordered
            pieces.
ACCESSION   AC123700
VERSION     AC123700.3  GI:28933837
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 200968)
  AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
  TITLE     Mus musculus, clone RP23-335A11
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 200968)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L.,
            Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
            Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzGerald,M., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I.,
            Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
            Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Lindblad-Toh,K.,
            Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N.,
            Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J.,
            Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
            Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
            Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C.,
            Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S.,
            Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N.,
```

```
            Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S.,
            Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J.,
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
            Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-JUN-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 200968)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
            Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
            Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
            Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K.,
            Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
            Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
            Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
            Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
            Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., MacLean,C.,
            Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
            Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
            Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
            Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
            Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,
            Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
            Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
            Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
            Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAR-2003) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Mar 13, 2003 this sequence version replaced gi:28201591.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L26431
                Center clone name: 335_A_11
            -------------- Summary Statistics
                Sequencing vector: Plasmid; n/a; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 198489 bases at least Q40
                Consensus quality: 199303 bases at least Q30
                Consensus quality: 199583 bases at least Q20
                Insert size: 197000; agarose-fp
                Insert size: 200168; sum-of-contigs
                Quality coverage: 9.9 in Q20 bases; agarose-fp
                Quality coverage: 9.7 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 9 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     44442: contig of 44442 bp in length
            *    44443     44542: gap of 100 bp
            *    44543     45239: contig of 697 bp in length
            *    45240     45339: gap of 100 bp
            *    45340     46467: contig of 1128 bp in length
            *    46468     46567: gap of 100 bp
            *    46568     55518: contig of 8951 bp in length
            *    55519     55618: gap of 100 bp
            *    55619    122673: contig of 67055 bp in length
            *   122674    122773: gap of 100 bp
            *   122774    137362: contig of 14589 bp in length
            *   137363    137462: gap of 100 bp
            *   137463    162118: contig of 24656 bp in length
            *   162119    162218: gap of 100 bp
            *   162219    187047: contig of 24829 bp in length
            *   187048    187147: gap of 100 bp
            *   187148    200968: contig of 13821 bp in length.
FEATURES             Location/Qualifiers
     source          1. .200968
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /clone="RP23-335A11"
                     /clone_lib="RPCI-23 Female Mouse BAC"
     misc_feature    1. .44442
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:left"
     misc_feature    44543. .45239
                     /note="assembly_fragment"
     misc_feature    45340. .46467
                     /note="assembly_fragment"
     misc_feature    46568. .55518
                     /note="assembly_fragment"
     misc_feature    55619. .122673
                     /note="assembly_fragment"
     misc_feature    122774. .137362
                     /note="assembly_fragment"
     misc_feature    137463. .162118
                     /note="assembly_fragment"
     misc_feature    162219. .187047
                     /note="assembly_fragment"
     misc_feature    187148. .200968
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
ORIGIN

Alignment Scores:
Pred. No.:              1.27e-158     Length:          200968
Score:                  2681.00       Matches:         771
Percent Similarity:     25.66%        Conservative:    28
Best Local Similarity:  24.76%        Mismatches:      42
Query Match:            56.01%        Indels:          2278
DB:                     2             Gaps:            18

US-09-970-944-2 (1-899) x AC123700 (1-200968)

Qy         59 LysProValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePhe--------- 75
              |||||     |||     :::      |||||      :::    |||
Db     101981 AAGCCATACTGCTTACATCCAAGAGATAAACCAGCAGCATCCCTGGAGTTCCCTCCAGGC 101922

Qy         76 ------LysCysAsnGly---------------GluTrpValArgGlnValAspHisVal 88
                    :::      |||                |||         :::::|||:::
Db     101921 CTCGGGAGGCTCACAGGGTCTATAAGCTGCTGTAGGTGGTGCGGGGGCATCAACCATTTG 101862

Qy         89 IleGluArgSerThrAsp---------GlySerSerGlyGluProThrMetGluValArg 105
                 :::      :::                  :::|||   ||||||||||||||||
Db     101861 GGACAGGGAGACTCAGTGTGGCCTCTGAACCCAGCAGGATTGCCAACCATGGAGGTCCGG 101802

Qy        106 IleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGluGluTyrTrpCysGln 125
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     101801 ATCAACGTATCAAGGCAGCAGGTCGAGAAAGTGTTTGGGCTGGAGGAGTACTGGTGCCAG 101742

Qy        126 CysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAlaTyrIleArgIleAla 145
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     101741 TGTGTGGCATGGAGCTCCTCAGGAACCACCAAAAGCCAGAAGGCCTACATCCGGATTGCC 101682

Qy        145 ------------------------------------------------------------ 145

Db     101681 TGTGAGTACTGGACCAGGCCTGTGGCATCCGGGCATGGTCACTAGAGGGTATCCCCTGGA 101622

Qy        145 ------------------------------------------------------------ 145
```

```
Db     101621 GAGAGGGCGGCAGGCTTCCTGGAGGGTAGAGGCAGTGCCTAGGAACATAAATCACCCCAC 101562

Qy        145 ------------------------------------------------------------ 145

Db     101561 ACTCCTCAATGCACATGGTGCCTCCTTCGGCCATGACGTGTGTCATGACGTCTGCAGGGC 101502

Qy        146 ---------------------------------------ArgLeuArgLysAsnPheGlu 152
                                                     |||||||||||||||||||||
Db     101501 CAAGAGTCAGCTGATCCAGACCCCGGTTTTCCTCCTCCAGATTTGCGCAAGAACTTTGAG 101442

Qy        153 GlnGluProLeuAlaLysGluValSerLeuGluGlnGlyIleValLeuProCysArgPro 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     101441 CAGGAGCCGCTGGCCAAGGAAGTGTCACTGGAGCAAGGCATTGTGCTACCTTGTCGCCCC 101382

Qy        173 ProGluGlyIleProProAlaGluVal--------------------------------- 181
              |||||||||||||||||||||||||||
Db     101381 CCGGAAGGAATCCCCCCAGCTGAGGT-GAGTGGAGGGTGCGGGCTTCCCAGACCTCTATC 101323

Qy        181 ------------------------------------------------------------ 181

Db     101322 CTACCTAAAGCTCAGACTCCTGGCTGCCACCTCCCCAGTGGTCGCTCTATACCCTGTTCA 101263

Qy        182 ------------------------------------------------GluTrpLeuArg 185
                                                              ||||||||||||
Db     101262 GGCCAGCACGGGCCAGCCGTGTAGTGTCACGGCTTCTCCCCCTAGGTGGAGTGGCTCCGA 101203

Qy        186 AsnGluAspLeuValAspProSerLeuAspProAsnValTyrIleThrArgGluHisSer 205
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     101202 AATGAGGACCTCGTGGACCCCTCCCTCGACCCCAATGTGTACATCACACGGGAGCACAGC 101143

Qy        206 LeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThrCysValAlaLysAsn 225
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     101142 CTAGTCGTGCGGCAGGCCCGCCTGGCCGACACTGCCAACTACACCTGCGTGGCCAAGAAC 101083

Qy        226 IleValAlaArgArgArgSerAlaSerAlaAlaValIleValTyrVal------------ 241
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db     101082 ATCGTGGCCCGTCGCCGAAGCGCCTCTGCGGCCGTCATTGTTTATGG-TGGGTCCTGGGC 101024

Qy        241 ------------------------------------------------------------ 241

Db     101023 TCCCCAGTCACAGGAAGAGGTACTGAGGTGCCAACAGGGACTTGGTGTGTGGCTCAGAGG 100964

Qy        241 ------------------------------------------------------------ 241

Db     100963 ACGGAAAGGCATGGAGCAAGAACCAGGCTAACCAGACAGTGGCTTGACTAGGAAAGAGAG 100904

Qy        241 ------------------------------------------------------------ 241

Db     100903 CAGAAGTGTATCTGGGCAAAGGGATGGACATGTGGGTCACTTGTGAGTGGGGAGCATGG 100844

Qy        241 ------------------------------------------------------------ 241

Db     100843 TGCAAGAGACGAAGAGACAAGAGGAGGCTCAGAGGGACGTAACACATCAACCTTGGGCGT 100784

Qy        241 ------------------------------------------------------------ 241

Db     100783 CTCAGTCCCCGAAATGGGAGAGAGCGCCTCAGAATGAGACGCCTTCCTCCCCCAAGAGCA 100724

Qy        241 ------------------------------------------------------------ 241

Db     100723 GAAGGAGGTCGGATAGGAGCTGCCCTGCGAGAAGCTGGGTTCTAGCTGGAGGTGCCCTGT 100664

Qy        241 ------------------------------------------------------------ 241

Db     100663 GACCTGTAGTTTAGTCTGGCCACCCTGACACTTGCCAAGAGTAGAGTATCTGAAGGCTCT 100604

Qy        241 ------------------------------------------------------------ 241

Db     100603 GCCAGCCCCTGGGGCCCAGCAGCCTAGGCTTGGTCCTCCTCAGGGTTTCTCCAGGCACAG 100544

Qy        241 ------------------------------------------------------------ 241

Db     100543 GAGCAGCCCACTCACCCCGAGCTTATAAGAACTTTGCCCTCCTGCCCCATCCCTTGGCCC 100484

Qy        242 ---------------------------------------------------------Asn 242
                                                                       |||
Db     100483 AGCATGTCTGAGGAAGCCCCCGCTCCCTGACCCCAGTCCTTCTCTGTCCGGCCAGTGAAC 100424

Qy        243 GlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArgGlyTrp 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     100423 GGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGTGCCAGCTGTGGGCGTGGCTGG 100364

Qy        263 GlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPheCysGlu 282
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     100363 CAGAAACGGAGCCGGAGCTGCACCAACCCGGCACCTCTCAACGGGGGCGCCTTCTGTGAG 100304

Qy        283 GlyGlnAsnVal------------------------------------------------ 286
              ||||||||||||
Db     100303 GGGCAGAATGTCCAGAAAACAGCCTGCGCCACTCTGTGCCCAGGTACCGGCAGCCACCGA 100244

Qy        286 ------------------------------------------------------------ 286

Db     100243 CTCCTGCATCGTTCCTCACCACGCCATCTCCGTGGCCTGGCTCCATCGCACCCACCAGCC 100184

Qy        286 ------------------------------------------------------------ 286

Db     100183 TGCCCGCCCCCTCCCCCATGGCTCCATCCCACCCGCCTGCCACACACAGGGCCAGGCTCA 100124

Qy        286 ------------------------------------------------------------ 286

Db     100123 GTCTGAGGTGGGGATCAGGGTGCAGTGTGGTCAGAGTCAAGGGCTCGGTTTAGGGTCTAA 100064

Qy        286 ------------------------------------------------------------ 286

Db     100063 GGTGCAGTTTAGGACAGGATCTAGGCTCCGCCTGGGGCTGCACAATCCACCTGTGCCTG 100004

Qy        286 ------------------------------------------------------------ 286

Db     100003 GACCAGAAGGGCCATCATGTGACTAGGAGCCCTTACCGCCTCAGGACAGGGACAGCCTCC 99944

Qy        286 ------------------------------------------------------------ 286

Db      99943 CTGGGTTGAGAGAGCCGAGCAGGCCCCAGCATCATTGCGTCTCCTTGTCATTAGCATGTG 99884

Qy        287 ------------------HisAspArgThrValSerSer--------------------- 293
                                |||   |||:::      |||
Db      99883 TCATCAAGGCTAGCTGGACACTCAAGGTCCGACCGATCAGAGGGCCTGACCCTGGCATTG 99824

Qy        293 ------------------------------------------------------------ 293

Db      99823 TACCCAAGCCCAGGCCCTGCCTGGTCCAGATGACTAGCGGGGCCACCCAGGCCAGCCACC 99764

Qy        294 ---------------------------------------------LeuLeuValSerVal 298
                                                           |||     |||||
Db      99763 CAGGCCAACCTCCTAAGCTGACTGTCCCATTCCCTCTCTGCCTGTCTCTCTTGTTCAGTG 99704

Qy        299 AspGlySerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrp 318
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      99703 GATGGGAGCTGGAGCCCATGGAGTAAGTGGTCAGCCTGCGGGCTTGACTGCACCCACTGG 99644

Qy        319 ArgSerArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThr 338
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db      99643 CGGAGCCGGGAGTGCTCCGACCCAGCGCCCCGCAACGGAGGTGAGGAGTGCCGGGGTGCT 99584

Qy        339 AspLeuAspThrArgAsnCysThrSerAspLeuCysValHis------------------ 352
              ||||||||||||||||||||||||||||||||||||:::|||
Db      99583 GACCTGGACACCCGCAACTGTACCAGTGACCTCTGCCTGCACAGTGAGTCTTCTCTGCCC 99524

Qy        352 ------------------------------------------------------------ 352

Db      99523 CAGGGCTCACCTCCGTCTGACAGCCCTATGGGCTGTGGTCAGGGCAGCCACTCTCCTGCC 99464

Qy        352 ------------------------------------------------------------ 352

Db      99463 TGTGTCTCTCCAAAGCTTGGGAGCCCAGGGGTCCCTTTCATCTCTAGCACACTCTTCCCT 99404
```

```
Qy        352 ------------------------------------------------------------ 352
Db      99403 GTTACACAGACACACCACCTTCCCCCTGAGACACATGCAGATGTGAATCTTGACCCCTGT 99344
Qy        352 ------------------------------------------------------------ 352
Db      99343 ACATCCAACTGCTCCCCCGCGGCTTTGAAATTATCCCGTTCCCCAAGCTCACACATAAGT 99284
Qy        352 ------------------------------------------------------------ 352
Db      99283 CCTGTGGGGGTGCTGGGCTTGTGAGGCCTGTTAATTACAAGGCCATAACTAACAATTAAA 99224
Qy        352 ------------------------------------------------------------ 352
Db      99223 GATCTCTGTTGGATTGTTTTTTTCCTTTAGACTACACCCTCCCCTGTTCTTTCAATTAAA 99164
Qy        352 ------------------------------------------------------------ 352
Db      99163 AACCTAATTATCTGTGTCAATTCCTTTCTGATGGACATTTCTTCCATTATGCAAAACTCT 99104
Qy        352 ------------------------------------------------------------ 352
Db      99103 TAAAAATCGATAGTGCTTCCTCCAAGCAGTTGGTCCCCCACCCACCCCAGTTTCAGGGCT 99044
Qy        352 ------------------------------------------------------------ 352
Db      99043 GGTGGTATACACAAGGGCTCACACTGTCCTGGAAACACCTGAGAAACTGACCACCCATGT 98984
Qy        352 ------------------------------------------------------------ 352
Db      98983 ATCTGCCTATTTTCTATTTCTATAATGAACTAGCAAAAGAAATTTATTAGTTTTAGAGGC 98924
Qy        352 ------------------------------------------------------------ 352
Db      98923 TCAAATTCCAAATAGCGTGATGCTGGCTCTGGTAAATGTGCCTTCTACCCCAAGCCACAT 98864
Qy        352 ------------------------------------------------------------ 352
Db      98863 TTCTTGGTGGCAGATAGCACCATGTCAGTACTGTAGACAAGAGAGAGATCACGTGTCAGG 98804
Qy        352 ------------------------------------------------------------ 352
Db      98803 AAGAGAAAAAGACAGACACAGAGACAGGTAGACACGGAAAGATGGACAGCCCCATGGTCT 98744
Qy        352 ------------------------------------------------------------ 352
Db      98743 AAGAGCAGCACTCATAATGATGTAATGCCTCCCCACTGGGGCCCACCTCCTAAAAGGCCC 98684
Qy        352 ------------------------------------------------------------ 352
Db      98683 CACTGCTGCTCATATTGCCACAATGGGGACCACACCTGTCAATATGAATCCTTAGGCAGG 98624
Qy        352 ------------------------------------------------------------ 352
Db      98623 CATGTGCAAACTATAGCACCATACTTAGGGTCTCTCAGGAGTTCACCTTGCCCCAGACAC 98564
Qy        352 ------------------------------------------------------------ 352
Db      98563 CCTGACCCCCAGCCTAGGAGCCAAGGATATACCACCACTCACCCCGGTGAAGTAGCTCTT 98504
Qy        352 ------------------------------------------------------------ 352
Db      98503 CCCCTGCTCCACTCTGCCTGTGAGGGTGTGACTTGCTCAATCCCAAGCAAGGTTGTGGTG 98444
Qy        352 ------------------------------------------------------------ 352
Db      98443 GGTATAGAACTAGGATGTACCCACCAATACCACGCTCCCTGCCACGTGCCCACTGCCCAG 98384
Qy        352 ------------------------------------------------------------ 352
Db      98383 GGGGTGCTGAGATCCTATTCTCCTGGGGATAAACCATGGGCCTCTCCTCTGAAGCCTCTA 98324

Qy        353 ----------------------------------------Ser-AlaSerGlyProGluAspVa 360
                                                      |||  :::||||||||||||||||
Db      98323 CATCCCTGCCATTCTCAAGACCACACTTTCTCTTGCAGCCTCTTCCGGCCCCGAGGACGT 98264
Qy        360 lAlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuValLe 380
              |||||||||:::||||||:::||||||||||||||||||||:::||||||||||||||||
Db      98263 GGCTCTCTACATCGGCCTCGTCGCCGTGGCCGTGTGCCTCATCTTGCTGCTGCTGGTCCT 98204
Qy        380 uIleLeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIl 400
              |:::|||:::||||||||||||||||||||||||||||||||||||||||||||||||||
Db      98203 CGTCCTCATCTACTGCCGCAAGAAGGAAGGACTGGACTCAGACGTGGCTGACTCATCCAT 98144
Qy        400 eLeuThrSerGlyPheGlnProValSerIleLysProSerLysAla--------------- 415
              ||||||||||||||||||||||||||||||||||||||||||||||
Db      98143 CCTTACCTCAGGCTTCCAGCCTGTCAGCATCAAGCCCAGCAAAGCAGGTGAGGTGCCACC 98084
Qy        415 ------------------------------------------------------------ 415
Db      98083 TACCCCTGCTCTCCTGCCTGAGCACCCCCAAGGCAAGGGCTGTTGATGGGGGAGTCGCTA 98024
Qy        416 -------------------------AspAsnProHisLeuLeuThrIleGlnProAspLeuS 428
                                       |||||||||||||||||||||||||||||||||||||
Db      98023 CCTGAATCCATCTTTATCCTGTAGACAATCCCCATCTGCTCACCATCCAACCGGACCTCA 97964
Qy        428 er---ThrThrThrThrTyrGlnGlySerLeuCysProArgGlnAspGlyProSerProL 447
              ||   |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      97963 GCACCACCACGACCACCTACCAGGGCAGCCTGTGTCCCCGGCAGGATGGACCCAGCCCCA 97904
Qy        447 ysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeuGlyGlyGlyArgHisThrLeuH 467
              ||||||||||:::||||||||||||||||||||||||||    |||||||||||||||||
Db      97903 AGTTCCAGCTCTCTAATGGTCACCTGCTCAGCCCACTGGGCAGTGGCCGCCATACGCTGC 97844
Qy        467 isHisSerSerProThrSerGluAlaGluGluPheValSerArgLeuSerThrGlnAsnT 487
              |||||||||||||||||||||||||||:::||||||||||||||||||||||||||||||
Db      97843 ACCACAGCTCCCCCACCTCTGAGGCTGAGGACTTCGTCTCCCGCCTCTCCACCCAAAACT 97784
Qy        487 yrPheArgSerLeuProArgGlyThrSerAsnMetThrTyrGlyThrPheAsnPheLeuG 507
              |||||||||||||||||||||||||||||||||||||    |||||||||||||||||||
Db      97783 ACTTTCGTTCTCTGCCCCGCGGTACCAGCAACATGGCCTATGGGACCTTCAACTTCCTCG 97724
Qy        507 lyGlyArgLeuMetIleProAsn------------------------------------- 514
              ||||||||||||||||||||||||
Db      97723 GGGGCCGGCTGATGATCCCTAACACAGGTAGGCCCCTAGAACTCCCCAGAGAACATGAGA 97664
Qy        514 ------------------------------------------------------------ 514
Db      97663 GGTAAGACTAGTCTTGTCATCTTTTCGTCCCAGAGTGCACCAGCCTCGTGCAGAGTGGCC 97604
Qy        514 ------------------------------------------------------------ 514
Db      97603 CAGTCTGGCCAGTATAGGAAAGGATGCAGTTGGCGCTATAGCAAAGCTGTGGAAAGGGTC 97544
Qy        514 ------------------------------------------------------------ 514
Db      97543 TCCTTGTTGCTTGTGGACACGAAAGGATTGTATGTTCATTTATCACACAGTCAGCCTTAC 97484
Qy        514 ------------------------------------------------------------ 514
Db      97483 ATGTTGATAGGTGCTGCATCGGGCCCTAACACCAGACCCCAGAGGCTCACAAAACAAGAC 97424
Qy        514 ------------------------------------------------------------ 514
Db      97423 AGCAGAGCACCCCTTAGAACCTGCCCCCTCCCAGCGTCCTCCATCTCATCCACCACTCCA 97364
Qy        514 ------------------------------------------------------------ 514
Db      97363 CAGAAGTGCCTCAGACAGGGTCCCTGAGACCTCTTGTTAGAAGATCCACAGACCCTCCAA 97304
Qy        514 ------------------------------------------------------------ 514
Db      97303 AGGCCATCTTTGGAGAAATCTCTGCAACCTCTGACAAGGCTGACCACCCCACTCTGACAT 97244
Qy        514 ------------------------------------------------------------ 514
```

```
Db      97243 TCTCTGTTCCCTCTCCTTAGGTGACCCCACATACTCCTGTTCTTGCCCATCCCTGGCTGT 97184
Qy        514 ------------------------------------------------------------ 514
Db      97183 TCCGGTATAGGCGTTCCTCAGGGATACCCCTCCCCAACCCACCCCCACTTCCGCCCGTGC 97124
Qy        514 ------------------------------------------------------------ 514
Db      97123 CATTTCTGAGTGATCTCACCTACTCCTCCATCCCAGCCCCCAGTTCTTGCAGAGCATTGG 97064
Qy        514 ------------------------------------------------------------ 514
Db      97063 AACTCTGCATTCAGCTGCCTCTTGGACAGCTCCATCTGCCTGTCTCACAGCACCTCGGAG 97004
Qy        514 ------------------------------------------------------------ 514
Db      97003 TCTTTGTGCCGAAAACAGAACCGTCTCCCACCCTTTCTCTCAGCTACCTCCTCCCTGGCT 96944
Qy        514 ------------------------------------------------------------ 514
Db      96943 TTCCCCGGCTCACGGTTGGCACAGTGTCATTGCTTAAGCTAGAAACTGGCCTGGGGTCAC 96884
Qy        514 ------------------------------------------------------------ 514
Db      96883 CTTGGGCTCCTGCTCTTTCTATTCCCACCCCCTCGCTATAAGTCAGCTGTATCACATGAT 96824
Qy        514 ------------------------------------------------------------ 514
Db      96823 GTTGCATGCTCCCCCTTCCCAATAGCTCCTGTCTCCTCCTGCTCCTCATCCTGCCCTGGT 96764
Qy        514 ------------------------------------------------------------ 514
Db      96763 GCGTCACCACACCCAGGCTAGCACCTCCCTTACCTCAACCGCGGCATCTACACATTTACT 96704
Qy        514 ------------------------------------------------------------ 514
Db      96703 GAATGAATCTAGAGTTGATTATGCCCCCTCCCCACTCAATCCCTTCTAAGGTTCCCAGTT 96644
Qy        514 ------------------------------------------------------------ 514
Db      96643 GCTCTCTACAGAGCCCCAGTGACAAGGCTAATGCAACCCACCCACTGCTGGCTAGCCGCT 96584
Qy        514 ------------------------------------------------------------ 514
Db      96583 GTGTGCTTTTCCCTCCTATATGCAGTTCTGTATGTGACAGCCCCTACAGACCTTTGGCAC 96524
Qy        514 ------------------------------------------------------------ 514
Db      96523 CTGAACACATGACACTGTCCCTGGCTCATAAGTAATGGTCTTTCTACCTTTCCTCCTGGC 96464
Qy        514 ------------------------------------------------------------ 514
Db      96463 TCATCTGTTGTTGCTGCCTGTGCCTCATACACCCCTCAGGACAGGGCATGAGAAATTCTC 96404
Qy        514 ------------------------------------------------------------ 514
Db      96403 ATAGCTCAGTGAAGTGAGCTGGTCCCTGTTAGCTGTAACAAAAAGGCAGACAGATGGTGG 96344
Qy        514 ------------------------------------------------------------ 514
Db      96343 GTATTCCGTGGCCCCTGCAGTGTGCCCACCCCTGTGTCCCTCTCTGTTGTCTAGTTGCCT 96284
Qy        514 ------------------------------------------------------------ 514
Db      96283 CTCTCCTTGCTGGACTGAGAAAGCTAGCGCCCTTTTTCCTCTTAGCTTTCTGAGTGGCCA 96224
Qy        514 ------------------------------------------------------------ 514
Db      96223 GCATGAGTATTCAATCTTGAATACTACTCTCCAGCAGGCATTCACACCAGGAGGTGTGTG 96164
Qy        514 ------------------------------------------------------------ 514

Db      96163 ACTGCCCCCGACTCCAGGCAGCATCAGGTTCTACGGTCAGCACAGGTTAGTGAAGGAAGC 96104
Qy        514 ------------------------------------------------------------ 514
Db      96103 AAAAGATTCGAGCACATGGTTCGAGCTGTCTGTGAGACAGCTCGGTGGTAAGACCCATC 96044
Qy        514 ------------------------------------------------------------ 514
Db      96043 ACTCGGTGGTAAGACACGGACAGGTACACATCGTAGAGCACACTTAAACCACTCGACAGC 95984
Qy        514 ------------------------------------------------------------ 514
Db      95983 TTCAGTGTCGGGAGGCTGTGTGACCCTATGGAACCATCACTGTGAGTTTGTAGAATATAA 95924
Qy        514 ------------------------------------------------------------ 514
Db      95923 GTGTCCTTACATAGCCCTTGAACAGCCTGAGAGTCCAGCCCAGCAAACAGTGGAGCTATC 95864
Qy        514 ------------------------------------------------------------ 514
Db      95863 TGCTGTTTTAGGGGCAAAGTCACAGAATGTCTCCCCTTTTCCTGCCCTCAGAGCAGGACC 95804
Qy        514 ------------------------------------------------------------ 514
Db      95803 TCAATGCTGCTCTCCGTGATGTCTCAGCCCTGCCCTTGCTGTCTCCACTACAAAACTCCC 95744
Qy        514 ------------------------------------------------------------ 514
Db      95743 CTTCCTCCACACACTGATACCACTCCTGGGGCAGAGGCGCATCAACAGACCAGGAGGACA 95684
Qy        514 ------------------------------------------------------------ 514
Db      95683 TTGGAGCCGTAGATCTTGGCCCTTGAGGGCTATTATAGACCTGCCCCCCATCCCCGCAAG 95624
Qy        514 ------------------------------------------------------------ 514
Db      95623 GATGGAGACAATGACAGGAAGAGACCAGTTCTTAGAGGGGACAGGACTAAGGGGGGAGGA 95564
Qy        514 ------------------------------------------------------------ 514
Db      95563 GCTTGAGCAAGACCAAAGGCCCCTGTTCTGGAACCGCTGAACGTAGCAAAGTTGGGCAGA 95504
Qy        514 ------------------------------------------------------------ 514
Db      95503 GGGACCGACTGCTGGGATGACAGTCAAGAGCCTTCCATGGCAGGCCTGGAAGATGGGGAA 95444
Qy        514 ------------------------------------------------------------ 514
Db      95443 GGTTTAAGCAGAAAGTACCCTGGTCCCAAAGAACGGAACAGGATGTAGGACGCACTTCC 95384
Qy        514 ------------------------------------------------------------ 514
Db      95383 TTGGTGTCTTCTTGCCCTACTTTAGCTTCTAGCTGGGTTGCCTCAGATGTTCTTGGAGTC 95324
Qy        514 ------------------------------------------------------------ 514
Db      95323 TCCCTGACTCAGAGTCGTGCCTATGTGGGTGGTGGACAGTATGGCAGTCCTTGAAACACC 95264
Qy        514 ------------------------------------------------------------ 514
Db      95263 TCTGTCCCCAGACAAGGGCGTCTGAAGAGGAAACTCGTGGGTCTGGCTAGTTCCTGGGTA 95204
Qy        515 ---------------------------------------------ThrGlyIleSerLeuL 520
                                                           |||||||||||||||
Db      95203 GTTATTGGGTATCTGTAGCCACAAGCCGTGACCCCTCTGCTTTTACAGGAATCAGCCTCC 95144
Qy        520 euIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHisL 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      95143 TCATACCCCCGGACGCCATCCCCCGAGGAAAGATCTACGAGATCTACCTCACTCTGCACA 95084
Qy        540 ysProGluAspVal---------------------------------------------- 544
              ||||||||||||||
Db      95083 AGCCAGAAGACGTGAGGTGTGTGCCCTGCCTGCTGTGGTGGGGGGGGGGGGAGGGCA 95024
```

```
Qy        544 ------------------------------------------------------------ 544

Db      95023 CCGTCTGCTGCTCCTATCGCTGGAACCACCCCAATTCCCCGCCCAGCACTGCTGTACCTG 94964

Qy        544 ------------------------------------------------------------ 544

Db      94963 TTCTGACCTGGCCCGAGGGAGCACAGGGGATGAGGCGCACTTAAGCTCACCCTGATTTGC 94904

Qy        544 ------------------------------------------------------------ 544

Db      94903 AACTTCCGTCCTTACGGTGTCAACCACCCCACTGCAGAGTCTGTCTCCTAACCCCACGGT 94844

Qy        545 -----------------------------------------------------ArgLeuP 547
                                                                   |||||||
Db      94843 CTCTAGGACTGGCCACTGACGCCTTTCCCTCTCCACCCATGTTTCCCCACTTGAGGTTGC 94784

Qy        547 roLeuAlaGlyCysGlnThrLeuLeuSerProIleValSerCysGlyProProGlyValL 567
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      94783 CCCTAGCTGGCTGTCAGACCCTGCTGAGTCCTATCGTTAGCTGTGGGCCCCCAGGAGTCC 94724

Qy        567 euLeuThrArgProValIleLeuAlaMetAspHisCysGlyGluProSerProAspSerT 587
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      94723 TGCTCACCCGGCCAGTCATCCTTGCCATGGACCACTGCGGGGAGCCCAGTCCCGACAGCT 94664

Qy        587 rpSerLeuArgLeuLysLysGlnSerCysGluGlySerTrpGlu---------------- 601
              ||||||||||||||||||||||||||||||||||||||||||||
Db      94663 GGAGCCTGCGCCTCAAAAAGCAGTCCTGTGAGGGCAGCTGGGAGGTGAGTTGGGAAGACT 94604

Qy        601 ------------------------------------------------------------ 601

Db      94603 ACCTGAGGCCACCCAACCAAATCACTGGCCCCACACTGGGCATGGCCTCCCATTCCTTTG 94544

Qy        601 ------------------------------------------------------------ 601

Db      94543 GCAGCTATGTGGTCTCGAGTCTCCCCAAACTGAGCTGAGGGTGGCCCGGAGGCTCTAAGG 94484

Qy        602 ------------------------------------------GlnAspValLeuHisLeu 607
                                                        ||||||||||||||||||
Db      94483 ACTCCTAAGTTGGTTAGGCACTGACCCACAAGGCCCCGCCCACAGGACGTGCTGCACCTT 94424

Qy        608 GlyGluGluAlaProSerHisLeuTyrTyrCysGlnLeuGluAlaSerAlaCysTyrVal 627
              ||||||||:::|||||||||||||||||||||||||||||||||||   ||||||||||||
Db      94423 GGTGAGGAGTCGCCCTCTCATCTCTACTACTGCCAGCTGGAGGCCGGGGCCTGCTATGTC 94364

Qy        628 PheThrGluGlnLeuGlyArgPheAlaLeuValGlyGluAlaLeuSerValAlaAlaAla 647
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      94363 TTCACCGAGCAGCTAGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCACC 94304

Qy        648 LysArgLeuLysLeuLeuLeuPheAlaProValAlaCysThrSerLeuGluTyrAsnIle 667
              ||||||||:::|||||||||||||||||||||||||||||||||||||||||||||||||
Db      94303 AAGCGCCTCAGGCTCCTTCTGTTTGCCCCTGTGGCCTGTACGTCCCTCGAGTACAACATC 94244

Qy        668 ArgValTyrCysLeuHisAspThrHisAspAlaLeuLys--------------------- 680
              ||||||||||||||||||||||||||||||||||||||
Db      94243 CGAGTGTACTGCCTGCACGACACCCACGATGCTCTCAA-GGTACTGCCGCGCACAGCCCC 94185

Qy        680 ------------------------------------------------------------ 680

Db      94184 ATTGGCGGACCACCCAGCCCTTGGCATCCAACTGCCTTGTCTTGGGGGTGACAGGGAGGG 94125

Qy        681 ---------------------------------------------GluValValGlnLeu 685
                                                           |||||||||||||||
Db      94124 TTTGCTCCTACCGTCGGGCCTTGCTGACCACCCCGGGCACTGCAGGAGGTGGTGCAGCTG 94065

Qy        686 GluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgValLeuHisPheLysAspSer 705
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      94064 GAGAAGCAGCTGGGTGGACAGCTGATCCAGGAGCCCCGTGTCCTGCACTTCAAAGACAGT 94005

Qy        706 TyrHisAsnLeuArgLeuSerIleHisAspValProSerSerLeuTrpLysSerLysLeu 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      94004 TACCACAACCTACGTCTGTCCATCCACGACGTGCCCAGCTCCCTGTGGAAGAGCAAGCTC 93945

Qy        726 LeuValSerTyrGln--------------------------------------------- 730
              |||||||||||||||
Db      93944 CTTGTCAGCTACCA-GGTGAGACCCTGAGCCCACGACCAGAGCGCTCCAGGGGCTGCTGG 93886

Qy        731 ------------------------------------------GluIleProPheTyrHis 736
                                                        ||||||||||||||||||
Db      93885 CGCTGGGCACGCTTGACGCCTCCTCTTTGCCCCTCTCCATAGGAGATCCCTTTTTACCAC 93826

Qy        737 IleTrpAsnGlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerPro 756
              |||||||||||||||||||:::||||||||||||||||||||||||||||||||||:::
Db      93825 ATCTGGAATGGCACTCAGCAGTATCTGCACTGCACCTTCACCCTGGAGCGCGTCAATGCC 93766

Qy        757 SerThrSerAspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSer 776
              ||||||||||||||||||||||:::|||||||||||||||||||||||||||||||||||
Db      93765 AGCACCAGCGACCTGGCCTGCAAGGTGTGGGTGTGGCAGGTGGAGGGAGATGGACAGAGC 93706

Qy        777 PheSerIleAsnPheAsnIleThrLys--------------------------------- 785
              |||:::|||||||||||||||||||||
Db      93705 TTCAACATCAACTTTAACATCACTAAGGTGGGCGACAAGATGGTGTGCTCCTAGTACCGA 93646

Qy        785 ------------------------------------------------------------ 785

Db      93645 CCATCCCAGGCAGGCTCCACCTCTGCGCGTAGCTGCCCACCAGCCACCGGCCTAACTTGA 93586

Qy        785 ------------------------------------------------------------ 785

Db      93585 GCTGCCCCTTCTAGCTTCTTCCAAGTTTCTGGGGACATAGGTGTAGGTGCTATATGAGGA 93526

Qy        785 ------------------------------------------------------------ 785

Db      93525 GCGGGAGGAGGTAGAGTTGTCCTGGGGGCTCTGGAGCGCTTCCCCAAAGAGGGACTTTGC 93466

Qy        785 ------------------------------------------------------------ 785

Db      93465 ACCGGGCCCCAAAGGATAGGCAGGAGCGGGAATCTGACAGACTGCAGGGTTGGGAGAAGG 93406

Qy        785 ------------------------------------------------------------ 785

Db      93405 CATTTGGGGCAGCAGGCACTGAGCTGAGAGCATAGCCATGGATGGTCAGCAGAGCAATAG 93346

Qy        785 ------------------------------------------------------------ 785

Db      93345 TGTGGGGAGCAGGGAGGTGTTCTCCGTCACCAACTGTGGCCAGGTGACTTCATCTCTGAG 93286

Qy        785 ------------------------------------------------------------ 785

Db      93285 ATCTAAAGAGTAAAGCCCTGGTGTACCAGGGAGAGGATGCCCTGCTGCCCCCATTGTCTG 93226

Qy        785 ------------------------------------------------------------ 785

Db      93225 GGGTTATCTGGAAACTCCCCTGGCCGAGGCGGCAGGGGTCAGAGTTCTAACTGGTACCTG 93166

Qy        786 ----------------------AspThrArgPheAlaGluLeuLeuAlaLeuGluSerGl 798
                                    |||||||||||||||||:::||||||||||||||||||
Db      93165 TTGCAAACATCCCATGATACAGGACACGAGGTTTGCTGAAATGCTGGCTCTGGAGAGTGA 93106

Qy        798 uAlaGlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGl 818
              |  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      93105 AGGGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCA 93046

Qy        818 nLysIleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAl 838
              ||||||||:::|||||||||||||||||||||   ||||||||||||||||||||||||||
Db      93045 AAAGATCATTACCAGCCTGGACCCACCCTGCAGCCGGGGCGCCGACTGGCGAACTCTAGC 92986

Qy        838 aGlnLysLeuHisLeuAspSer-------------------------------------- 845
              ||||||||||||||||||||||
Db      92985 CCAGAAACTTCACCTGGACAG-GTGGGTGGAGCTAGACAGAGGGAGGCTGGTTGGGCAAG 92927

Qy        845 ------------------------------------------------------------ 845

Db      92926 CGAGGTAGACAGGAAGGAGGCCTGTAGGCTCTGTGCCACAGCCTCTAGTGGTGTATTTCA 92867

Qy        846 ----------------------------------------------------HisLeuSe 848
```

```
                                                                           ||||||||
Db      92866 GGGCAGCCTGGAGTGAGGCCAGACTACTGACAGACTGCCTCTTCTTTGCAGCCATCTTAG 92807

Qy        848 rPhePheAlaSerLysProSerProThrAlaMetIleLeuAsnLeuTrpGluAlaArgHi 868
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      92806 CTTCTTTGCCTCCAAGCCCAGCCCTACAGCCATGATCCTCAACCTATGGGAGGCGCGGCA 92747

Qy        868 sPheProAsnGlyAsnLeuSerGlnLeuAlaAlaAlaValAlaGlyLeuGlyGlnProAs 888
              ||||||||||||||||||    ||||||||||||||||||||||||||||||||||||||
Db      92746 CTTCCCCAACGGCAACCTCGGCCAGCTGGCCGCAGCTGTGGCCGGACTGGGCCAGCCAGA 92687

Qy        888 pAlaGlyLeuPheThrValSerGluAlaGluCys 899
              ||||||||||||||||||||||||||||||||||
Db      92686 TGCTGGCCTCTTCACCGTGTCAGAGGCCGAGTGC 92653

Search completed: October  6, 2004, 21:08:28
Job time : 9114 secs
```

```
OM protein - protein search, using sw model

Run on:          October  6, 2004, 18:17:48 ; Search time 100 Seconds
                                              (without alignments)
                                              2892.975 Million cell updates/sec

Title:           US-09-970-944-2
Perfect score:   4787
Sequence:        1 MAVRPGLWPALLGIVLAAWL..........AVAGLGQPDAGLFTVSEAEC 899

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        1351062 seqs, 321799191 residues

Total number of hits satisfying chosen parameters:      1351062

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Published_Applications_AA:*
                 1:  /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
                 2:  /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
                 3:  /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
                 4:  /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
                 5:  /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
                 6:  /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
                 7:  /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
                 8:  /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
                 9:  /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
                 10:  /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
                 11:  /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
                 12:  /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
                 13:  /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
                 14:  /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
                 15:  /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
                 16:  /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
                 17:  /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
                 18:  /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 4787 | 100.0 | 899 | 11 | US-09-970-944-2 | Sequence 2, Appli |
| 2 | 4698.5 | 98.2 | 898 | 12 | US-09-918-779-2 | Sequence 2, Appli |
| 3 | 4698.5 | 98.2 | 898 | 16 | US-10-624-932-2 | Sequence 2, Appli |
| 4 | 4545.5 | 95.0 | 898 | 10 | US-09-933-261-5 | Sequence 5, Appli |
| 5 | 4545.5 | 95.0 | 898 | 11 | US-09-970-944-13 | Sequence 13, Appl |
| 6 | 4545.5 | 95.0 | 898 | 14 | US-10-256-702-5 | Sequence 5, Appli |
| 7 | 4545.5 | 95.0 | 898 | 14 | US-10-240-154-16 | Sequence 16, Appl |
| 8 | 4379.5 | 91.5 | 842 | 16 | US-10-311-623-1 | Sequence 1, Appli |
| 9 | 2819 | 58.9 | 544 | 11 | US-09-970-944-14 | Sequence 14, Appl |
| 10 | 2789.5 | 58.3 | 557 | 10 | US-09-933-261-6 | Sequence 6, Appli |
| 11 | 2789.5 | 58.3 | 557 | 14 | US-10-256-702-6 | Sequence 6, Appli |
| 12 | 2705.5 | 56.5 | 931 | 11 | US-09-970-944-15 | Sequence 15, Appl |
| 13 | 2705.5 | 56.5 | 931 | 12 | US-10-087-684-35 | Sequence 35, Appl |
| 14 | 2705.5 | 56.5 | 931 | 12 | US-09-972-211-121 | Sequence 121, App |
| 15 | 2705.5 | 56.5 | 931 | 12 | US-10-037-417-117 | Sequence 117, App |
| 16 | 2705.5 | 56.5 | 931 | 12 | US-10-096-625-121 | Sequence 121, App |
| 17 | 2705.5 | 56.5 | 1010 | 12 | US-10-218-779-35 | Sequence 35, Appl |
| 18 | 2680.5 | 56.0 | 931 | 11 | US-09-970-944-16 | Sequence 16, Appl |
| 19 | 2680.5 | 56.0 | 931 | 12 | US-09-972-211-125 | Sequence 125, App |
| 20 | 2680.5 | 56.0 | 931 | 12 | US-10-096-625-125 | Sequence 125, App |
| 21 | 2677.5 | 55.9 | 931 | 12 | US-10-037-417-120 | Sequence 120, App |
| 22 | 2673.5 | 55.8 | 931 | 11 | US-09-970-944-17 | Sequence 17, Appl |
| 23 | 2673.5 | 55.8 | 931 | 12 | US-10-087-684-36 | Sequence 36, Appl |
| 24 | 2673.5 | 55.8 | 931 | 12 | US-10-218-779-36 | Sequence 36, Appl |
| 25 | 2673.5 | 55.8 | 931 | 12 | US-09-972-211-122 | Sequence 122, App |
| 26 | 2673.5 | 55.8 | 931 | 12 | US-10-037-417-118 | Sequence 118, App |
| 27 | 2673.5 | 55.8 | 931 | 12 | US-10-037-417-119 | Sequence 119, App |
| 28 | 2673.5 | 55.8 | 931 | 12 | US-10-096-625-122 | Sequence 122, App |
| 29 | 2497 | 52.2 | 945 | 12 | US-10-087-684-34 | Sequence 34, Appl |
| 30 | 2497 | 52.2 | 945 | 12 | US-10-218-779-34 | Sequence 34, Appl |
| 31 | 2497 | 52.2 | 945 | 12 | US-09-972-211-124 | Sequence 124, App |
| 32 | 2497 | 52.2 | 945 | 12 | US-10-037-417-121 | Sequence 121, App |
| 33 | 2497 | 52.2 | 945 | 12 | US-10-096-625-124 | Sequence 124, App |
| 34 | 2493 | 52.1 | 945 | 12 | US-10-087-684-33 | Sequence 33, Appl |
| 35 | 2493 | 52.1 | 945 | 12 | US-10-218-779-33 | Sequence 33, Appl |
| 36 | 2493 | 52.1 | 945 | 12 | US-09-972-211-123 | Sequence 123, App |
| 37 | 2493 | 52.1 | 945 | 12 | US-10-096-625-123 | Sequence 123, App |
| 38 | 2490.5 | 52.0 | 946 | 12 | US-09-972-211-56 | Sequence 56, Appl |
| 39 | 2490.5 | 52.0 | 946 | 12 | US-10-096-625-56 | Sequence 56, Appl |
| 40 | 2490 | 52.0 | 943 | 10 | US-09-933-261-7 | Sequence 7, Appli |
| 41 | 2490 | 52.0 | 943 | 14 | US-10-256-702-7 | Sequence 7, Appli |
| 42 | 2484 | 51.9 | 933 | 12 | US-10-087-684-2 | Sequence 2, Appli |
| 43 | 2484 | 51.9 | 933 | 12 | US-10-087-684-4 | Sequence 4, Appli |
| 44 | 2484 | 51.9 | 933 | 12 | US-10-218-779-2 | Sequence 2, Appli |
| 45 | 2484 | 51.9 | 933 | 12 | US-10-218-779-4 | Sequence 4, Appli |

ALIGNMENTS

```
RESULT 1
US-09-970-944-2
; Sequence 2, Application US/09970944
; Publication No. US20030204052A1
; GENERAL INFORMATION:
;  APPLICANT: Herrman, John L
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Shimkets, Richard A
;  TITLE OF INVENTION: No. US20030204052A1el Proteins and Nucleic Acids Encoding Same anc
;  TITLE OF INVENTION:  Antibodies Directed Against these Proteins
;  FILE REFERENCE: 21402-138
;  CURRENT APPLICATION NUMBER: US/09/970,944
;  CURRENT FILING DATE:  2002-05-02
;  PRIOR APPLICATION NUMBER: 60/237,862
;  PRIOR FILING DATE: 2000-10-04
;  NUMBER OF SEQ ID NOS: 62
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 899
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-970-944-2

  Query Match             100.0%;  Score 4787;  DB 11;  Length 899;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  899;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60

Qy         61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGEPTMEVRINVSRQQVEKVFGLE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGEPTMEVRINVSRQQVEKVFGLE 120

Qy        121 EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180
```

```
Qy        181 VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY 240

Qy        241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDG 300

Qy        301 SWSPWSKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCVHSASGPEDV 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 SWSPWSKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCVHSASGPEDV 360

Qy        361 ALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHL 420

Qy        421 LTIQPDLSTTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEEFVS 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 LTIQPDLSTTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEEFVS 480

Qy        481 RLSTQNYFRSLPRGTSNMTYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLHK 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 RLSTQNYFRSLPRGTSNMTYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLHK 540

Qy        541 PEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGSW 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 PEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGSW 600

Qy        601 EQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAAKRLKLLLFAPVAC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 EQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAAKRLKLLLFAPVAC 660

Qy        661 TSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSL 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSL 720

Qy        721 WKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSIN 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 WKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSIN 780

Qy        781 FNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQK 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 FNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQK 840

Qy        841 LHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 899
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 LHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 899
```

```
RESULT 2
US-09-918-779-2
; Sequence 2, Application US/09918779
; Publication No. US20030064369A1
; GENERAL INFORMATION:
;  APPLICANT: Taupier, Raymond
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Spaderna, Steven
;  APPLICANT:  Shimkets, Richard
;  APPLICANT:  Zerhusen, Bryan
;  APPLICANT:  Spytek, Kimberly
;  APPLICANT:  Shenoy, Suresh
;  APPLICANT:  Li, Li
;  APPLICANT:  Gusev, Vladimir
;  APPLICANT:  Grosse, William
;  APPLICANT:  Alsobrook, John
;  APPLICANT:  Lepley, Denise
;  APPLICANT:  Burgess, Catherine
;  APPLICANT:  Gerlach, Valerie
;  APPLICANT:  Ellerman, Karen
;  APPLICANT:  MacDougall, John
;  APPLICANT:  Stone, David
;  APPLICANT:  Smithson, Glennda
;  TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
;  FILE REFERENCE: 21402-074 US
;  CURRENT APPLICATION NUMBER: US/09/918,779
;  CURRENT FILING DATE:  2001-07-30
;  PRIOR APPLICATION NUMBER: 60/221,409
;  PRIOR FILING DATE: 2000-07-28
;  PRIOR APPLICATION NUMBER: 60/222,840
;  PRIOR FILING DATE: 2000-08-04
;  PRIOR APPLICATION NUMBER: 60/223,752
;  PRIOR FILING DATE: 2000-08-08
;  PRIOR APPLICATION NUMBER: 60/223,762
;  PRIOR FILING DATE: 2000-08-08
;  PRIOR APPLICATION NUMBER: 60/223,770
;  PRIOR FILING DATE: 2000-08-08
;  PRIOR APPLICATION NUMBER: 60/223,769
;  PRIOR FILING DATE: 2000-08-08
;  PRIOR APPLICATION NUMBER: 60/225,146
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/225,392
;  PRIOR FILING DATE: 2000-08-15
;  PRIOR APPLICATION NUMBER: 60/225,470
;  PRIOR FILING DATE: 2000-08-15
;  PRIOR APPLICATION NUMBER: 60/225,697
;  PRIOR FILING DATE: 2000-08-16
;  PRIOR APPLICATION NUMBER: 60/263,662
;  PRIOR FILING DATE: 2001-02-01
;  PRIOR APPLICATION NUMBER: 60/281,645
;  PRIOR FILING DATE: 2001-04-05
;  NUMBER OF SEQ ID NOS: 61
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 898
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-918-779-2

  Query Match             98.2%;  Score 4698.5;  DB 12;  Length 898;
  Best Local Similarity   98.7%;  Pred. No. 0;
  Matches  888;  Conservative    2;  Mismatches    7;  Indels    3;  Gaps    3;

Qy          1 MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60

Qy         61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGEPTMEVRINVSRQQVEKVFGLE 120
              ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db         61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSRQQVEKVFGLE 120

Qy        121 EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180

Qy        181 VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY 240

Qy        241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDG 300
              ||||||||||||||||||||||||||||||||||||||||||||||  :|  :  |  |||
Db        241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNV-QKTACATLCPVDG 299

Qy        301 SWSPWSKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCVHSASGPEDV 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        300 SWSPWSKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCVHSASGPEDV 359

Qy        361 ALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        360 ALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHL 419

Qy        421 LTIQPDLS-TTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEEFV 479
              |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        420 LTIQPDLSTTTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEEFV 479
```

```
Qy        480 SRLSTQNYFRSLPRGTSNMTYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLH 539
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 SRLSTQNYFRSLPRGTSNMTYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLH 539

Qy        540 KPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGS 599
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 KPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGS 599

Qy        600 WEQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAAKRLKLLLFAPVA 659
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 WE-DVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAAKRLKLLLFAPVA 658

Qy        660 CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSS 719
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        659 CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSS 718

Qy        720 LWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSI 779
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        719 LWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSI 778

Qy        780 NFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQ 839
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        779 NFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQ 838

Qy        840 KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 899
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        839 KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 898
```

```
RESULT 3
US-10-624-932-2
; Sequence 2, Application US/10624932
; Publication No. US20040096877A1
; GENERAL INFORMATION:
;  APPLICANT: Taupier, Raymond
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Spaderna, Steven
;  APPLICANT:  Shimkets, Richard
;  APPLICANT:  Zerhusen, Bryan
;  APPLICANT:  Spytek, Kimberly
;  APPLICANT:  Shenoy, Suresh
;  APPLICANT:  Li, Li
;  APPLICANT:  Gusev, Vladimir
;  APPLICANT:  Grosse, William
;  APPLICANT:  Alsobrook, John
;  APPLICANT:  Lepley, Denise
;  APPLICANT:  Burgess, Catherine
;  APPLICANT:  Gerlach, Valerie
;  APPLICANT:  Ellerman, Karen
;  APPLICANT:  MacDougall, John
;  APPLICANT:  Stone, David
;  APPLICANT:  Smithson, Glennda
;  TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
;  FILE REFERENCE: 21402-074 US
;  CURRENT APPLICATION NUMBER: US/10/624,932
;  CURRENT FILING DATE:  2003-07-21
;  PRIOR APPLICATION NUMBER: 09/918,779
;  PRIOR FILING DATE: 2001-07-03
;  PRIOR APPLICATION NUMBER: 60/221,409
;  PRIOR FILING DATE: 2000-07-28
;  PRIOR APPLICATION NUMBER: 60/222,840
;  PRIOR FILING DATE: 2000-08-04
;  PRIOR APPLICATION NUMBER: 60/223,752
;  PRIOR FILING DATE: 2000-08-08
;  PRIOR APPLICATION NUMBER: 60/223,762
;  PRIOR FILING DATE: 2000-08-08
;  PRIOR APPLICATION NUMBER: 60/223,770
;  PRIOR FILING DATE: 2000-08-08
;  PRIOR APPLICATION NUMBER: 60/223,769
;  PRIOR FILING DATE: 2000-08-08
;  PRIOR APPLICATION NUMBER: 60/225,146
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/225,392
;  PRIOR FILING DATE: 2000-08-15
;  PRIOR APPLICATION NUMBER: 60/225,470
;  PRIOR FILING DATE: 2000-08-15
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 61
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 898
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-624-932-2

  Query Match             98.2%;  Score 4698.5;  DB 16;  Length 898;
  Best Local Similarity   98.7%;  Pred. No. 0;
  Matches  888;  Conservative    2;  Mismatches    7;  Indels    3;  Gaps    3;

Qy          1 MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60

Qy         61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGEPTMEVRINVSRQQVEKVFGLE 120
              |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db         61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSRQQVEKVFGLE 120

Qy        121 EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180

Qy        181 VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY 240

Qy        241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDG 300
              |||||||||||||||||||||||||||||||||||||||||||     :|   :  |  |||
Db        241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNV-QKTACATLCPVDG 299

Qy        301 SWSPWSKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCVHSASGPEDV 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        300 SWSPWSKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCVHSASGPEDV 359

Qy        361 ALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        360 ALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHL 419

Qy        421 LTIQPDLS-TTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEEFV 479
              |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        420 LTIQPDLSTTTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEEFV 479

Qy        480 SRLSTQNYFRSLPRGTSNMTYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLH 539
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 SRLSTQNYFRSLPRGTSNMTYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLH 539

Qy        540 KPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGS 599
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 KPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGS 599

Qy        600 WEQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAAKRLKLLLFAPVA 659
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 WE-DVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAAKRLKLLLFAPVA 658

Qy        660 CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSS 719
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        659 CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSS 718

Qy        720 LWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSI 779
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        719 LWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSI 778

Qy        780 NFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQ 839
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        779 NFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQ 838
```

```
Qy        840 KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 899
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        839 KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 898


RESULT 4
US-09-933-261-5
; Sequence 5, Application US/09933261
; Publication No. US20030040046A1
;  GENERAL INFORMATION:
;       APPLICANT: Tessier-Lavigne, Marc
;                  Leonardo, E. David
;                  Hink, Lindsay
;                  Masu, Masayuki
;                  Kazuko, Keino-Masu
;       TITLE OF INVENTION: Netrin Receptors
;       NUMBER OF SEQUENCES: 8
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
;            STREET: 268 BUSH STREET, SUITE 3200
;            CITY: SAN FRANCISCO
;            STATE: CALIFORNIA
;            COUNTRY: USA
;            ZIP: 94104
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/933,261
;            FILING DATE: 20-Aug-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 08/808,982
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: OSMAN, RICHARD A
;            REGISTRATION NUMBER: 36,627
;            REFERENCE/DOCKET NUMBER: UC96-217
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (415) 343-4341
;            TELEFAX: (415) 343-4342
;   INFORMATION FOR SEQ ID NO: 5:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 898 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: No. US20030040046A1 Relevant
;            TOPOLOGY: No. US20030040046A1 Relevant
;       MOLECULE TYPE: peptide
;       SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-933-261-5

  Query Match             95.0%;  Score 4545.5;  DB 10;  Length 898;
  Best Local Similarity   94.7%;  Pred. No. 0;
  Matches  852;  Conservative   19;  Mismatches   26;  Indels    3;  Gaps    3;

Qy          1 MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60
              ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAVRPGLWPVLLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60

Qy         61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGEPTMEVRINVSRQQVEKVFGLE 120
              |||||||||||||||||||||||||||||||||| ||| |||||||||||||||||||||
Db         61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDSSSGLPTMEVRINVSRQQVEKVFGLE 120

Qy        121 EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180
              ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        121 EYWCQCVAWSSSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180

Qy        181 VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        181 VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSTSAAVIVY 240

Qy        241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDG 300
              ||||||||||||||||||||||||||||||||||||||||||||||  :|  : |  |||
Db        241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNV-QKTACATLCPVDG 299

Qy        301 SWSPWSKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCVHSASGPEDV 360
              ||| |||||||||||||||||||||||||||||||:| ||||||||||||:|:|| ||||
Db        300 SWSSWSKWSACGLDCTHWRSRECSDPAPRNGGEECRGADLDTRNCTSDLCLHTASCPEDV 359

Qy        361 ALYVGLIAVAVCLVLLLVLILVYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHL 420
              |||:||:|||||| ||| | |:||||||||||||||||||||||||||||||||||||||
Db        360 ALYIGLVAVAVCLFLLLALGLIYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHL 419

Qy        421 LTIQPDLS-TTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEEFV 479
              |||||||| |||||||||| |||||||||||:||||||||| |||||||||||||||:||
Db        420 LTIQPDLSTTTTTYQGSLCSRQDGPSPKFQLSNGHLLSPLGSGRHTLHHSSPTSEAEDFV 479

Qy        480 SRLSTQNYFRSLPRGTSNMTYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLH 539
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        480 SRLSTQNYFRSLPRGTSNMAYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLH 539

Qy        540 KPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGS 599
              ||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||
Db        540 KPEDVRLPLAGCQTLLSPVVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGS 599

Qy        600 WEQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAAKRLKLLLFAPVA 659
              || ||||||||:||||||||||| ||||||||||||||||||||||| |||:||||||||
Db        600 WE-DVLHLGEESPSHLYYCQLEAGACYVFTEQLGRFALVGEALSVAATKRLRLLLFAPVA 658

Qy        660 CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSS 719
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        659 CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSS 718

Qy        720 LWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSI 779
              |||||||||||||||||||||||:||||||||||:: ||||||||:||||||||||||:|
Db        719 LWKSKLLVSYQEIPFYHIWNGTQQYLHCTFTLERINASTSDLACKVWVWQVEGDGQSFNI 778

Qy        780 NFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQ 839
              ||||||||||||||||||| ||||||||||||||||||||||:|||||| ||||||||||
Db        779 NFNITKDTRFAELLALESEGGVPALVGPSAFKIPFLIRQKIIASLDPPCSRGADWRTLAQ 838

Qy        840 KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 899
              ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        839 KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLGQLAAAVAGLGQPDAGLFTVSEAEC 898


RESULT 5
US-09-970-944-13
; Sequence 13, Application US/09970944
; Publication No. US20030204052A1
; GENERAL INFORMATION:
;  APPLICANT: Herrman, John L
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Shimkets, Richard A
;  TITLE OF INVENTION: No. US20030204052A1el Proteins and Nucleic Acids Encoding Same and
;  TITLE OF INVENTION:  Antibodies Directed Against these Proteins
;  FILE REFERENCE: 21402-138
;  CURRENT APPLICATION NUMBER: US/09/970,944
;  CURRENT FILING DATE:  2002-05-02
;  PRIOR APPLICATION NUMBER: 60/237,862
;  PRIOR FILING DATE: 2000-10-04
;  NUMBER OF SEQ ID NOS: 62
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
;   LENGTH: 898
;   TYPE: PRT
;   ORGANISM: Rattus norvegicus
US-09-970-944-13

  Query Match             95.0%;  Score 4545.5;  DB 11;  Length 898;
  Best Local Similarity   94.7%;  Pred. No. 0;
  Matches  852;  Conservative   19;  Mismatches   26;  Indels    3;  Gaps    3;

Qy          1 MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60
```

```
             ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 MAVRPGLWPVLLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60

Qy        61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGEPTMEVRINVSRQQVEKVFGLE 120
             |||||||||||||||||||||||||||||||||| ||| |||||||||||||||||||||
Db        61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDSSSGLPTMEVRINVSRQQVEKVFGLE 120

Qy       121 EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180
             ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db       121 EYWCQCVAWSSSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180

Qy       181 VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY 240
             |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db       181 VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSTSAAVIVY 240

Qy       241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDG 300
             ||||||||||||||||||||||||||||||||||||||||||||||  :|  : |  |||
Db       241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNV-QKTACATLCPVDG 299

Qy       301 SWSPWSKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCVHSASGPEDV 360
             ||| |||||||||||||||||||||||||||||||:| ||||||||||||:|:|| ||||
Db       300 SWSSWSKWSACGLDCTHWRSRECSDPAPRNGGEECRGADLDTRNCTSDLCLHTASCPEDV 359

Qy       361 ALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHL 420
             |||:||:|||||| |||| | |:|||||||||||||||||||||||||||||||||||||
Db       360 ALYIGLVAVAVCLFLLLLALGLIYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHL 419

Qy       421 LTIQPDLS-TTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEEFV 479
             |||||||| |||||||||| |||||||||||:||||||||| |||||||||||||||:||
Db       420 LTIQPDLSTTTTTYQGSLCSRQDGPSPKFQLSNGHLLSPLGSGRHTLHHSSPTSEAEDFV 479

Qy       480 SRLSTQNYFRSLPRGTSNMTYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLH 539
             ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       480 SRLSTQNYFRSLPRGTSNMAYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLH 539

Qy       540 KPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGS 599
             ||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||
Db       540 KPEDVRLPLAGCQTLLSPVVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGS 599

Qy       600 WEQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAAKRLKLLLFAPVA 659
             || ||||||||:||||||||||| ||||||||||||||||||||||| |||:||||||||
Db       600 WE-DVLHLGEESPSHLYYCQLEAGACYVFTEQLGRFALVGEALSVAATKRLRLLLFAPVA 658

Qy       660 CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSS 719
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       659 CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSS 718

Qy       720 LWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSI 779
             |||||||||||||||||||||||:||||||||||:: ||||||||:||||||||||||:|
Db       719 LWKSKLLVSYQEIPFYHIWNGTQQYLHCTFTLERINASTSDLACKVWVWQVEGDGQSFNI 778

Qy       780 NFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQ 839
             ||||||||||||||||||| |||||||||||||||||||||:||||||| ||||||||||
Db       779 NFNITKDTRFAELLALESEGGVPALVGPSAFKIPFLIRQKIIASLDPPCSRGADWRTLAQ 838

Qy       840 KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 899
             ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db       839 KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLGQLAAAVAGLGQPDAGLFTVSEAEC 898
```

```
RESULT 6
US-10-256-702-5
; Sequence 5, Application US/10256702
; Publication No. US20030059859A1
;  GENERAL INFORMATION:
;       APPLICANT: Tessier-Lavigne, Marc
;                  Leonardo, E. David
;                  Hink, Lindsay
;                  Masu, Masayuki
;                  Kazuko, Keino-Masu
;       TITLE OF INVENTION: Netrin Receptors
;       NUMBER OF SEQUENCES: 8
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
;            STREET: 268 BUSH STREET, SUITE 3200
;            CITY: SAN FRANCISCO
;            STATE: CALIFORNIA
;            COUNTRY: USA
;            ZIP: 94104
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/256,702
;            FILING DATE: 27-Sep-2002
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/09/933,261
;            FILING DATE: 20-Aug-2001
;            APPLICATION NUMBER: 08/808,982
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: OSMAN, RICHARD A
;            REGISTRATION NUMBER: 36,627
;            REFERENCE/DOCKET NUMBER: UC96-217
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (415) 343-4341
;            TELEFAX: (415) 343-4342
;  INFORMATION FOR SEQ ID NO: 5:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 898 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: No. US20030059859A1 Relevant
;            TOPOLOGY: No. US20030059859A1 Relevant
;       MOLECULE TYPE: peptide
;       SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-10-256-702-5

  Query Match             95.0%;  Score 4545.5;  DB 14;  Length 898;
  Best Local Similarity   94.7%;  Pred. No. 0;
  Matches  852;  Conservative   19;  Mismatches   26;  Indels    3;  Gaps    3;

Qy         1 MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60
             ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 MAVRPGLWPVLLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60

Qy        61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGEPTMEVRINVSRQQVEKVFGLE 120
             |||||||||||||||||||||||||||||||||| ||| |||||||||||||||||||||
Db        61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDSSSGLPTMEVRINVSRQQVEKVFGLE 120

Qy       121 EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180
             ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db       121 EYWCQCVAWSSSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180

Qy       181 VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY 240
             |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db       181 VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSTSAAVIVY 240

Qy       241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDG 300
             ||||||||||||||||||||||||||||||||||||||||||||||  :|  : |  |||
Db       241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNV-QKTACATLCPVDG 299

Qy       301 SWSPWSKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCVHSASGPEDV 360
             ||| |||||||||||||||||||||||||||||||:| ||||||||||||:|:|| ||||
Db       300 SWSSWSKWSACGLDCTHWRSRECSDPAPRNGGEECRGADLDTRNCTSDLCLHTASCPEDV 359

Qy       361 ALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHL 420
             |||:||:|||||| |||| | |:|||||||||||||||||||||||||||||||||||||
Db       360 ALYIGLVAVAVCLFLLLLALGLIYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHL 419

Qy       421 LTIQPDLS-TTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEEFV 479
             |||||||| |||||||||| |||||||||||:||||||||| |||||||||||||||:||
Db       420 LTIQPDLSTTTTTYQGSLCSRQDGPSPKFQLSNGHLLSPLGSGRHTLHHSSPTSEAEDFV 479
```

```
Qy      480 SRLSTQNYFRSLPRGTSNMTYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLH 539
Db      480 SRLSTQNYFRSLPRGTSNMAYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLH 539

Qy      540 KPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGS 599
Db      540 KPEDVRLPLAGCQTLLSPVVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGS 599

Qy      600 WEQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAAKRLKLLLFAPVA 659
Db      600 WE-DVLHLGEESPSHLYYCQLEAGACYVFTEQLGRFALVGEALSVAATKRLRLLLFAPVA 658

Qy      660 CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSS 719
Db      659 CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSS 718

Qy      720 LWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSI 779
Db      719 LWKSKLLVSYQEIPFYHIWNGTQQYLHCTFTLERINASTSDLACKVWVWQVEGDGQSFNI 778

Qy      780 NFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQ 839
Db      779 NFNITKDTRFAELLALESEGGVPALVGPSAFKIPFLIRQKIIASLDPPCSRGADWRTLAQ 838

Qy      840 KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 899
Db      839 KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLGQLAAAVAGLGQPDAGLFTVSEAEC 898
```

```
RESULT 7
US-10-240-154-16
; Sequence 16, Application US/10240154
; Publication No. US20030175741A1
; GENERAL INFORMATION:
;  APPLICANT: Cochran et al.
;  TITLE OF INVENTION: SCHIZOPHRENIA RELATED GENES
;  FILE REFERENCE: CKFW-P01-006
;  CURRENT APPLICATION NUMBER: US/10/240,154
;  CURRENT FILING DATE:  2001-04-02
;  PRIOR APPLICATION NUMBER: PCT/GB01/01486
;  PRIOR FILING DATE: 2001-04-02
;  NUMBER OF SEQ ID NOS: 34
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 16
;   LENGTH: 898
;   TYPE: PRT
;   ORGANISM: Rattus sp.
US-10-240-154-16

  Query Match             95.0%;  Score 4545.5;  DB 14;  Length 898;
  Best Local Similarity   94.7%;  Pred. No. 0;
  Matches  852;  Conservative   19;  Mismatches   26;  Indels    3;  Gaps    3;

Qy        1 MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60
Db        1 MAVRPGLWPVLLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60

Qy       61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGEPTMEVRINVSRQQVEKVFGLE 120
Db       61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDSSSGLPTMEVRINVSRQQVEKVFGLE 120

Qy      121 EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180
Db      121 EYWCQCVAWSSSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180

Qy      181 VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY 240
Db      181 VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSTSAAVIVY 240

Qy      241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDG 300
Db      241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNV-QKTACATLCPVDG 299

Qy      301 SWSPWSKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCVHSASGPEDV 360
Db      300 SWSSWSKWSACGLDCTHWRSRECSDPAPRNGGEECRGADLDTRNCTSDLCLHTASCPEDV 359

Qy      361 ALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHL 420
Db      360 ALYIGLVAVAVCLFLLLLALGLIYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHL 419

Qy      421 LTIQPDLS-TTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEEFV 479
Db      420 LTIQPDLSTTTTTYQGSLCSRQDGPSPKFQLSNGHLLSPLGSGRHTLHHSSPTSEAEDFV 479

Qy      480 SRLSTQNYFRSLPRGTSNMTYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLH 539
Db      480 SRLSTQNYFRSLPRGTSNMAYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLH 539

Qy      540 KPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGS 599
Db      540 KPEDVRLPLAGCQTLLSPVVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGS 599

Qy      600 WEQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAAKRLKLLLFAPVA 659
Db      600 WE-DVLHLGEESPSHLYYCQLEAGACYVFTEQLGRFALVGEALSVAATKRLRLLLFAPVA 658

Qy      660 CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSS 719
Db      659 CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSS 718

Qy      720 LWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSI 779
Db      719 LWKSKLLVSYQEIPFYHIWNGTQQYLHCTFTLERINASTSDLACKVWVWQVEGDGQSFNI 778

Qy      780 NFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQ 839
Db      779 NFNITKDTRFAELLALESEGGVPALVGPSAFKIPFLIRQKIIASLDPPCSRGADWRTLAQ 838

Qy      840 KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 899
Db      839 KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLGQLAAAVAGLGQPDAGLFTVSEAEC 898
```

```
RESULT 8
US-10-311-623-1
; Sequence 1, Application US/10311623
; Publication No. US20040023244A1
; GENERAL INFORMATION:
;  APPLICANT: INCYTE GENOMICS, INC.; GRIFFIN, Jennifer A.
;  APPLICANT:  KALLICK, Deborah A.; TRIBOULEY, Catherine M.
;  APPLICANT:  YUE, Henry; NGUYEN, Danniel B.
;  APPLICANT:  TANG, Y. Tom; LAL, Preeti G.
;  APPLICANT:  POLICKY, Jennifer L.; AZIMZAI, Yalda
;  APPLICANT:  LU, Dyung Aina M.; GRAUL, Richard C.
;  APPLICANT:  YAO, Monique G.; BURFORD, Neil
;  APPLICANT:  HAFALIA, April J. A.; BAUGHN, Mariah R.
;  APPLICANT:  BANDMAN, Olga; ARVIZU, Chandra S.
;  APPLICANT:  YANG, Junming; XU, Yuming
;  APPLICANT:  GANDHI, Ameena R.; WARREN, Bridget A.
;  APPLICANT:  DING, Li; SANJANWALA, Madhusudan M.
;  APPLICANT:  DUGGAN, Brendan M.; LU, Yan
;  TITLE OF INVENTION: RECEPTORS
;  FILE REFERENCE: PF-0793 USN
;  CURRENT APPLICATION NUMBER: US/10/311,623
;  CURRENT FILING DATE:  2002-12-17
;  PRIOR APPLICATION NUMBER: US 01/19942
;  PRIOR FILING DATE: 2001-06-21
;  PRIOR APPLICATION NUMBER: US 60/214,027
;  PRIOR FILING DATE: 2000-06-21
;  PRIOR APPLICATION NUMBER: US 60/228,045
;  PRIOR FILING DATE: 2000-08-25
;  PRIOR APPLICATION NUMBER: US 60/255,104
;  PRIOR FILING DATE: 2000-12-12
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PERL Program
```

```
; SEQ ID NO 1
;   LENGTH: 842
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20040023244A1 6052371CD1
US-10-311-623-1

  Query Match             91.5%;  Score 4379.5;  DB 16;  Length 842;
  Best Local Similarity   93.1%;  Pred. No. 0;
  Matches  838;  Conservative    1;  Mismatches    2;  Indels   59;  Gaps    3;

Qy          1 MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60

Qy         61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGEPTMEVRINVSRQQVEKVFGLE 120
              |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db         61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSRQQVEKVFGLE 120

Qy        121 EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180
              |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db        121 EYWCQCVAWSSSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180

Qy        181 VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY 240

Qy        241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDG 300
                                                                       |||
Db        241 ---------------------------------------------------------VDG 243

Qy        301 SWSPWSKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCVHSASGPEDV 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db        244 SWSPWSKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCVHTASGPEDV 303

Qy        361 ALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 ALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHL 363

Qy        421 LTIQPDLS-TTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEEFV 479
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 LTIQPDLSTTTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEEFV 423

Qy        480 SRLSTQNYFRSLPRGTSNMTYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLH 539
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 SRLSTQNYFRSLPRGTSNMTYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLH 483

Qy        540 KPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGS 599
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 KPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGS 543

Qy        600 WEQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAAKRLKLLLFAPVA 659
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        544 WE-DVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAAKRLKLLLFAPVA 602

Qy        660 CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSS 719
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        603 CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSS 662

Qy        720 LWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSI 779
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        663 LWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSI 722

Qy        780 NFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQ 839
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        723 NFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQ 782

Qy        840 KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 899
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        783 KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 842


RESULT 9
US-09-970-944-14
; Sequence 14, Application US/09970944
; Publication No. US20030204052A1
; GENERAL INFORMATION:
;  APPLICANT: Herrman, John L
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Shimkets, Richard A
;  TITLE OF INVENTION: No. US20030204052A1el Proteins and Nucleic Acids Encoding Same anc
;  TITLE OF INVENTION:  Antibodies Directed Against these Proteins
;  FILE REFERENCE: 21402-138
;  CURRENT APPLICATION NUMBER: US/09/970,944
;  CURRENT FILING DATE:  2002-05-02
;  PRIOR APPLICATION NUMBER: 60/237,862
;  PRIOR FILING DATE: 2000-10-04
;  NUMBER OF SEQ ID NOS: 62
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
;   LENGTH: 544
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-970-944-14

  Query Match             58.9%;  Score 2819;  DB 11;  Length 544;
  Best Local Similarity   99.6%;  Pred. No. 4.2e-238;
  Matches  540;  Conservative    0;  Mismatches    0;  Indels    2;  Gaps    2;

Qy        359 DVALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNP 418
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4 DVALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNP 63

Qy        419 HLLTIQPDLS-TTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEE 477
              |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db         64 HLLTIQPDLSTTTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEE 123

Qy        478 FVSRLSTQNYFRSLPRGTSNMTYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLT 537
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 FVSRLSTQNYFRSLPRGTSNMTYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLT 183

Qy        538 LHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCE 597
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 LHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCE 243

Qy        598 GSWEQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAAKRLKLLLFAP 657
              |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 GSWE-DVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAAKRLKLLLFAP 302

Qy        658 VACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVP 717
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 VACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVP 362

Qy        718 SSLWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSF 777
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 SSLWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSF 422

Qy        778 SINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTL 837
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 SINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTL 482

Qy        838 AQKLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEA 897
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 AQKLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEA 542

Qy        898 EC 899
              ||
Db        543 EC 544


RESULT 10
US-09-933-261-6
; Sequence 6, Application US/09933261
; Publication No. US20030040046A1
```

```
;   GENERAL INFORMATION:
;        APPLICANT: Tessier-Lavigne, Marc
;                   Leonardo, E. David
;                   Hink, Lindsay
;                   Masu, Masayuki
;                   Kazuko, Keino-Masu
;        TITLE OF INVENTION: Netrin Receptors
;        NUMBER OF SEQUENCES: 8
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
;             STREET: 268 BUSH STREET, SUITE 3200
;             CITY: SAN FRANCISCO
;             STATE: CALIFORNIA
;             COUNTRY: USA
;             ZIP: 94104
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/933,261
;             FILING DATE: 20-Aug-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/808,982
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: OSMAN, RICHARD A
;             REGISTRATION NUMBER: 36,627
;             REFERENCE/DOCKET NUMBER: UC96-217
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 343-4341
;             TELEFAX: (415) 343-4342
;   INFORMATION FOR SEQ ID NO: 6:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 557 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: No. US20030040046A1 Relevant
;             TOPOLOGY: No. US20030040046A1 Relevant
;        MOLECULE TYPE: peptide
;        SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-933-261-6

  Query Match             58.3%;  Score 2789.5;  DB 10;  Length 557;
  Best Local Similarity   96.4%;  Pred. No. 1.7e-235;
  Matches  538;  Conservative    2;  Mismatches   15;  Indels    3;  Gaps    3;

Qy        344 NCTSDLCVHSASGPEDVALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTS 403
              ||||| ||:|||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 NCTSDLXVHTASGPEDVALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTS 60

Qy        404 GFQPVSIKPSKADNPHLLTIQPDLS-TTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGG 462
              ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db         61 GFQPVSIKPSKADNPHLLTIQPDLSTTTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGG 120

Qy        463 RHTLHHSSPTSEAEEFVSRLSTQNYFRSLPRGTSNMTYGTFNFLGGRLMIPNTGISLLIP 522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RHTLHHSSPTSEAEEFVSRLSTQNYFRSLPRGTSNMTYGTFNFLGGRLMIPNTGISLLIP 180

Qy        523 PDAIPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEP 582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PDAIPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEP 240

Qy        583 SPDSWSLRLKKQSCEGSWEQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEAL 642
              |||||| |||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        241 SPDSWSLALKKQSCEGSWE-DVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEAL 299

Qy        643 SVAAAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHF 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db        300 SVAAAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHL 359

Qy        703 KDSYHNLRLSIHDVPSSLWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLA 762
              ||||| || |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        360 XDSYHNLXLSXHDVPSSLWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLA 419

Qy        763 CKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIIS 822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        420 CKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIIS 479

Qy        823 SLDPPCRRGADWRTLAQKLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVA 882
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 SLDPPCRRGADWRTLAQKLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVA 539

Qy        883 GLGQPDAGLFT-VSEAEC 899
              |       |:  |||||
Db        540 GTXPAGRWLLSQCSEAEC 557


RESULT 11
US-10-256-702-6
; Sequence 6, Application US/10256702
; Publication No. US20030059859A1
;   GENERAL INFORMATION:
;        APPLICANT: Tessier-Lavigne, Marc
;                   Leonardo, E. David
;                   Hink, Lindsay
;                   Masu, Masayuki
;                   Kazuko, Keino-Masu
;        TITLE OF INVENTION: Netrin Receptors
;        NUMBER OF SEQUENCES: 8
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
;             STREET: 268 BUSH STREET, SUITE 3200
;             CITY: SAN FRANCISCO
;             STATE: CALIFORNIA
;             COUNTRY: USA
;             ZIP: 94104
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/256,702
;             FILING DATE: 27-Sep-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/933,261
;             FILING DATE: 20-Aug-2001
;             APPLICATION NUMBER: 08/808,982
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: OSMAN, RICHARD A
;             REGISTRATION NUMBER: 36,627
;             REFERENCE/DOCKET NUMBER: UC96-217
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 343-4341
;             TELEFAX: (415) 343-4342
;   INFORMATION FOR SEQ ID NO: 6:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 557 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: No. US20030059859A1 Relevant
;             TOPOLOGY: No. US20030059859A1 Relevant
;        MOLECULE TYPE: peptide
;        SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-256-702-6

  Query Match             58.3%;  Score 2789.5;  DB 14;  Length 557;
  Best Local Similarity   96.4%;  Pred. No. 1.7e-235;
  Matches  538;  Conservative    2;  Mismatches   15;  Indels    3;  Gaps    3;

Qy        344 NCTSDLCVHSASGPEDVALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTS 403
              ||||| ||:|||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 NCTSDLXVHTASGPEDVALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTS 60
```

```
Qy        404 GFQPVSIKPSKADNPHLLTIQPDLS-TTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGG 462
Db         61 GFQPVSIKPSKADNPHLLTIQPDLSTTTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGG 120

Qy        463 RHTLHHSSPTSEAEEFVSRLSTQNYFRSLPRGTSNMTYGTFNFLGGRLMIPNTGISLLIP 522
Db        121 RHTLHHSSPTSEAEEFVSRLSTQNYFRSLPRGTSNMTYGTFNFLGGRLMIPNTGISLLIP 180

Qy        523 PDAIPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEP 582
Db        181 PDAIPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEP 240

Qy        583 SPDSWSLRLKKQSCEGSWEQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEAL 642
Db        241 SPDSWSLALKKQSCEGSWE-DVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEAL 299

Qy        643 SVAAAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHF 702
Db        300 SVAAAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHL 359

Qy        703 KDSYHNLRLSIHDVPSSLWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLA 762
Db        360 XDSYHNLXLSXHDVPSSLWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLA 419

Qy        763 CKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIIS 822
Db        420 CKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIIS 479

Qy        823 SLDPPCRRGADWRTLAQKLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVA 882
Db        480 SLDPPCRRGADWRTLAQKLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVA 539

Qy        883 GLGQPDAGLFT-VSEAEC 899
Db        540 GTXPAGRWLLSQCSEAEC 557
```

RESULT 12
US-09-970-944-15
; Sequence 15, Application US/09970944
; Publication No. US20030204052A1
; GENERAL INFORMATION:
; APPLICANT: Herrman, John L
; APPLICANT: Rastelli, Luca
; APPLICANT: Shimkets, Richard A
; TITLE OF INVENTION: No. US20030204052A1el Proteins and Nucleic Acids Encoding Same an
; TITLE OF INVENTION: Antibodies Directed Against these Proteins
; FILE REFERENCE: 21402-138
; CURRENT APPLICATION NUMBER: US/09/970,944
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: 60/237,862
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 931
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-970-944-15

```
  Query Match             56.5%;  Score 2705.5;  DB 11;  Length 931;
  Best Local Similarity   56.4%;  Pred. No. 9e-228;
  Matches  515;  Conservative  153;  Mismatches  214;  Indels   31;  Gaps   12;

Qy          9 PALLGIVLAAWLRGSGAQQSA---TVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVC 65
Db         26 PAL--ALLSASGTGSAAQDDEFFHELPETFPSDPPEPLPHFLIEPEEAYIVKNKPVNLYC 83

Qy         66 KAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGEPTMEVRINVSRQQVEKVFGLEEYWCQ 125
Db         84 KASPATQIYFKCNSEWVHQKDHVVDERVDETSGLIVREVSIEISRQQVEELFGPEDYWCQ 143

Qy        126 CVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAEVEWLR 185
Db        144 CVAWSSAGTTKSRKAYVRIAYLRKTFEQEPLGKEVSLEQEVLLQCRPPEGIPVAEVEWLK 203

Qy        186 NEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNGGW 245
Db        204 NEDIIDPAEDRNFYITIDHNLIIKQARLSDTANYTCVAKNIVAKRKSTTATVIVYVNGGW 263

Qy        246 STWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDGSWSPW 305
Db        264 STWTEWSVCNSRCGRGYQKRTRTCTNPAPLNGGAFCEGQSV-QKIACTTLCPVDGRWTSW 322

Qy        306 SKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCVHSASGPEDVALYVG 365
Db        323 SKWSTCGTECTHWRRRECTAPAPKNGGKDCDGLVLQSKNCTDGLCMQAAPDSDDVALYVG 382

Qy        366 L-IAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHLLTIQ 424
Db        383 IVIAVTVCLAITVVVALFVYRKNHRDFESDIIDSSALNGGFQPVNIKAARQD---LLAVP 439

Qy        425 PDL-STTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSS----PTSEAEEFV 479
Db        440 PDLTSAAAMYRGPVYALHD-VSDKIPMTNSPILDPLPNLKIKVYNSSGAVTPQDDLAEFS 498

Qy        480 SRLS---TQNYF---------RSLPRGT--SNMTYGTFNFLGGRLMIPNTGISLLIPPDA 525
Db        499 SKLSPQMTQSLLENEALNLKNQSLARQTDPSCTAFGTFNSLGGHLIIPNSGVSLLIPAGA 558

Qy        526 IPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPD 585
Db        559 IPQGRVYEMYVTVHRKENMRPPMEDSQTLLTPVVSCGPPGALLTRPVILTLHHCADPSTE 618

Qy        586 SWSLRLKKQSCEGSWEQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA 645
Db        619 DWKIQLKNQAVQGQWE-DVVVVGEENFTTPCYIQLDAEACHILTENLSTYALVGQSTTKA 677

Qy        646 AAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDS 705
Db        678 AAKRLKLAIFGPLCCSSLEYSIRVYCLDDTQDALKEVLQLERQMGGQLLEEPKALHFKGS 737

Qy        706 YHNLRLSIHDVPSSLWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKL 765
Db        738 IHNLRLSIHDIAHSLWKSKLLAKYQEIPFYHIWSGSQRNLHCTFTLERLSLNTVELVCKL 797

Qy        766 WVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLD 825
Db        798 CVRQVEGEGQIFQLNCTVSEEPTGIDLPLLDPASTITTVTGPSAFSIPLPIRQKLCSSLD 857

Qy        826 PPCRRGADWRTLAQKLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLG 885
Db        858 APQTRGHDWRMLAHKLNLDRYLNYFATKSSPTGVILDLWEAQNFPDGNLSMLAAVLEEMG 917

Qy        886 QPDAGLFTVSEAE 898
Db        918 RHETVVSLAAEGQ 930
```

RESULT 13
US-10-087-684-35
; Sequence 35, Application US/10087684
; Publication No. US20040029116A1
; GENERAL INFORMATION:
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: MacDougall, John R.
; APPLICANT: Millet, Isabelle
; APPLICANT: Ellerman, Karen
; APPLICANT: Stone, David J.
; APPLICANT: Grosse, William M.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Cathereine E.
; APPLICANT: Casman, Stacie, J.
; APPLICANT: Spytek, Kimberly A.

```
;  APPLICANT:  Boldog, Ferenc L.
;  APPLICANT:  Li, Li
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Mishra, Vishnu
;  APPLICANT:  Shenoy, Suresh G.
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Tchernev, Velizar T.
;  APPLICANT:  Vernet, Corine A.M.
;  APPLICANT:  Zerhusen, Bryan D.
;  APPLICANT:  Malyankar, Uriel M.
;  APPLICANT:  Guo, Xiaojia
;  APPLICANT:  Miller, Charles E.
;  APPLICANT:  Gangolli, Esha A.
;  TITLE OF INVENTION: PROTEINS AND NUCLEIC ACIDS ENCODING SAME
;  FILE REFERENCE: 21402-214 CIP
;  CURRENT APPLICATION NUMBER: US/10/087,684
;  CURRENT FILING DATE:  2003-03-10
;  PRIOR APPLICATION NUMBER: 60/253,834
;  PRIOR FILING DATE: 2000-11-29
;  PRIOR APPLICATION NUMBER: 60/250,926
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: 60/264,180
;  PRIOR FILING DATE: 2001-01-25
;  PRIOR APPLICATION NUMBER: 60/274,194
;  PRIOR FILING DATE: 2001-03-08
;  PRIOR APPLICATION NUMBER: 60/313,656
;  PRIOR FILING DATE: 2001-08-20
;  PRIOR APPLICATION NUMBER: 60/327,456
;  PRIOR FILING DATE: 2001-10-05
;  NUMBER OF SEQ ID NOS: 220
;  SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 35
;   LENGTH: 931
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-087-684-35

  Query Match             56.5%;  Score 2705.5;  DB 12;  Length 931;
  Best Local Similarity   56.4%;  Pred. No. 9e-228;
  Matches  515;  Conservative  153;  Mismatches  214;  Indels   31;  Gaps   12;

Qy          9 PALLGIVLAAWLRGSGAQQSA---TVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVC 65
              |||   :|:|    || ||         :     |     |: ||||:|||: |||||||| | |
Db         26 PAL--ALLSASGTGSAAQDDEFFHELPETFPSDPPEPLPHFLIEPEEAYIVKNKPVNLYC 83

Qy         66 KAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGEPTMEVRINVSRQQVEKVFGLEEYWCQ 125
              || |||||:||||  |||  |  |||::     | :||      || |  :|||||||::|| |:||||
Db         84 KASPATQIYFKCNSEWVHQKDHVVDERVDETSGLIVREVSIEISRQQVEELFGPEDYWCQ 143

Qy        126 CVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAEVEWLR 185
              |||||:||||:||:|||  |||   ||||||  ||||||  ::| ||||||| ||||||:
Db        144 CVAWSSAGTTKSRKAYVRIAYLRKTFEQEPLGKEVSLEQEVLLQCRPPEGIPVAEVEWLK 203

Qy        186 NEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNGGW 245
              |||::||:  |  |  |||  :|:|::::||||:||||||||||||||||:|:|  :| |||||||||
Db        204 NEDIIDPAEDRNFYITIDHNLIIKQARLSDTANYTCVAKNIVAKRKSTTATVIVYVNGGW 263

Qy        246 STWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDGSWSPW 305
              ||||||||::  ||||:|||:|:||||||||||||||||:|      :      :   |  |||  |:  |
Db        264 STWTEWSVCNSRCGRGYQKRTRTCTNPAPLNGGAFCEGQSV-QKIACTTLCPVDGRWTSW 322

Qy        306 SKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCVHSASGPEDVALYVG 365
              |||| ||  :|||||  |||:  |||:|||::|  |    |   ::|||    |||:  :|       :||||||
Db        323 SKWSTCGTECTHWRRRECTAPAPKNGGKDCDGLVLQSKNCTDGLCMQAAPDSDDVALYVG 382

Qy        366 L-IAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHLLTIQ 424
              :  |||  |||   :  ::|  :  ||   :          :||:  |||  |    ||||:||  ::  |       ||   :
Db        383 IVIAVTVCLAITVVVALFVYRKNHRDFESDIIDSSALNGGFQPVNIKAARQD---LLAVP 439

Qy        425 PDL-STTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSS----PTSEAEEFV 479
              |||  |        |:|  :        |   |   |  :||   :|  ||         :  :::||       |     ||
Db        440 PDLTSAAAMYRGPVYALHD-VSDKIPMTNSPILDPLPNLKIKVYNSSGAVTPQDDLAEFS 498

Qy        480 SRLS---TQNYF---------RSLPRGT--SNMTYGTFNFLGGRLMIPNTGISLLIPPDA 525
              |:||    ||:                  :||  |  |     :||||  |||  |:||||:|:||||     |
Db        499 SKLSPQMTQSLLENEALNLKNQSLARQTDPSCTAFGTFNSLGGHLIIPNSGVSLLIPAGA 558

Qy        526 IPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPD 585
              ||:|::||:|:|:|:  |::|  |:      ||||:|:||||||| |||||||   :  ||  :|| :
Db        559 IPQGRVYEMYVTVHRKENMRPPMEDSQTLLTPVVSCGPPGALLTRPVILTLHHCADPSTE 618

Qy        586 SWSLRLKKQSCEGSWEQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA 645
              | ::||  |: :|  ||  ||: :|||   :     | ||:|  ||::  || |    :||||:: :  |
Db        619 DWKIQLKNQAVQGQWE-DVVVVGEENFTTPCYIQLDAEACHILTENLSTYALVGQSTTKA 677

Qy        646 AAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDS 705
              |||||| :|  |:  |:||||:||||||  ||  |||||:|||:|:|||::||: ||||  |
Db        678 AAKRLKLAIFGPLCCSSLEYSIRVYCLDDTQDALKEVLQLERQMGGQLLEEPKALHFKGS 737

Qy        706 YHNLRLSIHDVPSSLWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKL 765
              |||||||||:   ||||||||   ||||||||||:|:||  ||||||||||:|  :|  :|  |||
Db        738 IHNLRLSIHDIAHSLWKSKLLAKYQEIPFYHIWSGSQRNLHCTFTLERLSLNTVELVCKL 797

Qy        766 WVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLD 825
              |  ||||:||  |  :|    ::::      :|    |:   :   :    :  |||||  ||    ||||:  ||||
Db        798 CVRQVEGEGQIFQLNCTVSEEPTGIDLPLLDPASTITTVTGPSAFSIPLPIRQKLCSSLD 857

Qy        826 PPCRRGADWRTLAQKLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLG 885
              |   ||  |||  ||  ||:||   :|::|:||  |||  :||:||||::||:||||  |||    :   :|
Db        858 APQTRGHDWRMLAHKLNLDRYLNYFATKSSPTGVILDLWEAQNFPDGNLSMLAAVLEEMG 917

Qy        886 QPDAGLFTVSEAE 898
               :  :   :  :| :
Db        918 RHETVVSLAAEGQ 930

RESULT 14
US-09-972-211-121
; Sequence 121, Application US/09972211
; Publication No. US20040048245A1
; GENERAL INFORMATION:
;  APPLICANT: Shimkets, Richard A
;  APPLICANT:  Taupier Jr, Raymond J
;  APPLICANT:  Burgess, Catherine E
;  APPLICANT:  Zerhusen, Bryan D
;  APPLICANT:  Mezes, Peter S
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Malyankar, Uriel M
;  APPLICANT:  Grosse, William M
;  APPLICANT:  Alsobrook II, John P
;  APPLICANT:  Lepley, Denise M
;  APPLICANT:  Spytek, Kimberly Ann
;  APPLICANT:  Li, Li
;  APPLICANT:  Edinger, Shlomit
;  APPLICANT:  Gerlach, Valerie
;  APPLICANT:  Ellerman, Karen
;  APPLICANT:  MacDougall, John R
;  APPLICANT:  Gunther, Erik
;  APPLICANT:  Millet, Isabelle
;  APPLICANT:  Stone, David J
;  APPLICANT:  Smithson, Glennda
;  APPLICANT:  Szekeres Jr, Edward S
;  TITLE OF INVENTION: No. US20040048245A1el Human Proteins, Polynucleotides Encoding The
;  TITLE OF INVENTION:  Methods Of Using The Same
;  FILE REFERENCE: 21402-141
;  CURRENT APPLICATION NUMBER: US/09/972,211
;  CURRENT FILING DATE:  2001-10-05
;  PRIOR APPLICATION NUMBER: 60/238,325
;  PRIOR FILING DATE: 2000-10-05
;  PRIOR APPLICATION NUMBER: 60/238,323
;  PRIOR FILING DATE: 2000-10-05
;  PRIOR APPLICATION NUMBER: 60/238,400
;  PRIOR FILING DATE: 2000-10-06
;  PRIOR APPLICATION NUMBER: 60/238,397
;  PRIOR FILING DATE: 2000-10-06
;  PRIOR APPLICATION NUMBER: 60/238,401
```

```
;  PRIOR FILING DATE: 2000-10-06
;  PRIOR APPLICATION NUMBER: 60/238,379
;  PRIOR FILING DATE: 2000-10-06
;  PRIOR APPLICATION NUMBER: 60/238,402
;  PRIOR FILING DATE: 2000-10-06
;  PRIOR APPLICATION NUMBER: 30/238,384
;  PRIOR FILING DATE: 2000-10-06
;  PRIOR APPLICATION NUMBER: 60/238,373
;  PRIOR FILING DATE: 2000-10-06
;  PRIOR APPLICATION NUMBER: 60/238,372
;  PRIOR FILING DATE: 2000-10-06
;  PRIOR APPLICATION NUMBER: 60/238,383
;  PRIOR FILING DATE: 2000-10-06
;  PRIOR APPLICATION NUMBER: 60/238,382
;  PRIOR FILING DATE: 2000-10-06
;  PRIOR APPLICATION NUMBER: 60/275,892
;  PRIOR FILING DATE: 2001-03-14
;  PRIOR APPLICATION NUMBER: 60/296,860
;  PRIOR FILING DATE: 2001-06-08
;  NUMBER OF SEQ ID NOS: 198
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 121
;   LENGTH: 931
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-972-211-121

  Query Match             56.5%;  Score 2705.5;  DB 12;  Length 931;
  Best Local Similarity   56.4%;  Pred. No. 9e-228;
  Matches  515;  Conservative  153;  Mismatches  214;  Indels   31;  Gaps   12;

Qy          9 PALLGIVLAAWLRGSGAQQSA---TVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVC 65
Db         26 PAL--ALLSASGTGSAAQDDEFFHELPETFPSDPPEPLPHFLIEPEEAYIVKNKPVNLYC 83

Qy         66 KAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGEPTMEVRINVSRQQVEKVFGLEEYWCQ 125
Db         84 KASPATQIYFKCNSEWVHQKDHVVDERVDETSGLIVREVSIEISRQQVEELFGPEDYWCQ 143

Qy        126 CVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAEVEWLR 185
Db        144 CVAWSSAGTTKSRKAYVRIAYLRKTFEQEPLGKEVSLEQEVLLQCRPPEGIPVAEVEWLK 203

Qy        186 NEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNGGW 245
Db        204 NEDIIDPAEDRNFYITIDHNLIIKQARLSDTANYTCVAKNIVAKRKSTTATVIVYVNGGW 263

Qy        246 STWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDGSWSPW 305
Db        264 STWTEWSVCNSRCGRGYQKRTRTCTNPAPLNGGAFCEGQSV-QKIACTTLCPVDGRWTSW 322

Qy        306 SKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCVHSASGPEDVALYVG 365
Db        323 SKWSTCGTECTHWRRRECTAPAPKNGGKDCDGLVLQSKNCTDGLCMQAAPDSDDVALYVG 382

Qy        366 L-IAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHLLTIQ 424
Db        383 IVIAVTVCLAITVVVALFVYRKNHRDFESDIIDSSALNGGFQPVNIKAARQD---LLAVP 439

Qy        425 PDL-STTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSS----PTSEAEEFV 479
Db        440 PDLTSAAAMYRGPVYALHD-VSDKIPMTNSPILDPLPNLKIKVYNSSGAVTPQDDLAEFS 498

Qy        480 SRLS---TQNYF---------RSLPRGT--SNMTYGTFNFLGGRLMIPNTGISLLIPPDA 525
Db        499 SKLSPQMTQSLLENEALNLKNQSLARQTDPSCTAFGTFNSLGGHLIIPNSGVSLLIPAGA 558

Qy        526 IPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPD 585
Db        559 IPQGRVYEMYVTVHRKENMRPPMEDSQTLLTPVVSCGPPGALLTRPVILTLHHCADPSTE 618

Qy        586 SWSLRLKKQSCEGSWEQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA 645
Db        619 DWKIQLKNQAVQGQWE-DVVVVGEENFTTPCYIQLDAEACHILTENLSTYALVGQSTTKA 677

Qy        646 AAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDS 705
Db        678 AAKRLKLAIFGPLCCSSLEYSIRVYCLDDTQDALKEVLQLERQMGGQLLEEPKALHFKGS 737

Qy        706 YHNLRLSIHDVPSSLWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKL 765
Db        738 IHNLRLSIHDIAHSLWKSKLLAKYQEIPFYHIWSGSQRNLHCTFTLERLSLNTVELVCKL 797

Qy        766 WVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLD 825
Db        798 CVRQVEGEGQIFQLNCTVSEEPTGIDLPLLDPASTITTVTGPSAFSIPLPIRQKLCSSLD 857

Qy        826 PPCRRGADWRTLAQKLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLG 885
Db        858 APQTRGHDWRMLAHKLNLDRYLNYFATKSSPTGVILDLWEAQNFPDGNLSMLAAVLEEMG 917

Qy        886 QPDAGLFTVSEAE 898
Db        918 RHETVVSLAAEGQ 930

RESULT 15
US-10-037-417-117
; Sequence 117, Application US/10037417
; Publication No. US20040052806A1
; GENERAL INFORMATION:
;  APPLICANT: Kekuda, Ramesh
;  APPLICANT:  Alsobrook II, John P
;  APPLICANT:  Tchernev, Velizar T
;  APPLICANT:  Liu, Xiaohong
;  APPLICANT:  Spytek, Kimberly A
;  APPLICANT:  Patturajan, Meera
;  APPLICANT:  Grosse, William M
;  APPLICANT:  Lepley, Denise M
;  APPLICANT:  Burgess, Catherine E
;  APPLICANT:  Vernet, Corine A.M.
;  APPLICANT:  Li, Li
;  APPLICANT:  Gorman, Linda
;  APPLICANT:  Edinger, Shlomit R
;  APPLICANT:  Sciore, Paul
;  APPLICANT:  Ellerman, Karen
;  APPLICANT:  Malyankar, Uriel M
;  APPLICANT:  Rothenberg, Mark
;  APPLICANT:  Stone, David J
;  APPLICANT:  Boldog, Ferenc L
;  APPLICANT:  Guo, Xiaojia
;  APPLICANT:  Shenoy, Suresh G
;  APPLICANT:  Anderson, David W
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Taupier Jr, Raymond J
;  APPLICANT:  Miller, Charles E
;  APPLICANT:  Eisen, Andrew J
;  TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
;  FILE REFERENCE: 21402-235
;  CURRENT APPLICATION NUMBER: US/10/037,417
;  CURRENT FILING DATE:  2002-09-20
;  PRIOR APPLICATION NUMBER: 60/260,018
;  PRIOR FILING DATE: 2001-01-05
;  PRIOR APPLICATION NUMBER: 60/260,360
;  PRIOR FILING DATE: 2001-01-08
;  PRIOR APPLICATION NUMBER: 60/272,411
;  PRIOR FILING DATE: 2001-02-28
;  PRIOR APPLICATION NUMBER: 60/272,817
;  PRIOR FILING DATE: 2001-03-02
;  PRIOR APPLICATION NUMBER: 60/291,186
;  PRIOR FILING DATE: 2001-05-15
;  PRIOR APPLICATION NUMBER: 60/303,231
;  PRIOR FILING DATE: 2001-07-05
;  PRIOR APPLICATION NUMBER: 60/305,060
;  PRIOR FILING DATE: 2001-07-12
;  PRIOR APPLICATION NUMBER: 60/318,405
```

```
;  PRIOR FILING DATE: 2001-09-10
;  PRIOR APPLICATION NUMBER: 60/318,700
;  PRIOR FILING DATE: 2001-09-12
;  NUMBER OF SEQ ID NOS: 227
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 117
;   LENGTH: 931
;   TYPE: PRT
;   ORGANISM: Caenorhabditis elegans
US-10-037-417-117

  Query Match             56.5%;  Score 2705.5;  DB 12;  Length 931;
  Best Local Similarity   56.4%;  Pred. No. 9e-228;
  Matches  515;  Conservative  153;  Mismatches  214;  Indels   31;  Gaps   12;

Qy          9 PALLGIVLAAWLRGSGAQQSA---TVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVC 65
Db         26 PAL--ALLSASGTGSAAQDDEFFHELPETFPSDPPEPLPHFLIEPEEAYIVKNKPVNLYC 83

Qy         66 KAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGEPTMEVRINVSRQQVEKVFGLEEYWCQ 125
Db         84 KASPATQIYFKCNSEWVHQKDHVVDERVDETSGLIVREVSIEISRQQVEELFGPEDYWCQ 143

Qy        126 CVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAEVEWLR 185
Db        144 CVAWSSAGTTKSRKAYVRIAYLRKTFEQEPLGKEVSLEQEVLLQCRPPEGIPVAEVEWLK 203

Qy        186 NEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNGGW 245
Db        204 NEDIIDPAEDRNFYITIDHNLIIKQARLSDTANYTCVAKNIVAKRKSTTATVIVYVNGGW 263

Qy        246 STWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDGSWSPW 305
Db        264 STWTEWSVCNSRCGRGYQKRTRTCTNPAPLNGGAFCEGQSV-QKIACTTLCPVDGRWTSW 322

Qy        306 SKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCVHSASGPEDVALYVG 365
Db        323 SKWSTCGTECTHWRRRECTAPAPKNGGKDCDGLVLQSKNCTDGLCMQAAPDSDDVALYVG 382

Qy        366 L-IAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHLLTIQ 424
Db        383 IVIAVTVCLAITVVVALFVYRKNHRDFESDIIDSSALNGGFQPVNIKAARQD---LLAVP 439

Qy        425 PDL-STTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSS----PTSEAEEFV 479
Db        440 PDLTSAAAMYRGPVYALHD-VSDKIPMTNSPILDPLPNLKIKVYNSSGAVTPQDDLAEFS 498

Qy        480 SRLS---TQNYF---------RSLPRGT--SNMTYGTFNFLGGRLMIPNTGISLLIPPDA 525
Db        499 SKLSPQMTQSLLENEALNLKNQSLARQTDPSCTAFGTFNSLGGHLIIPNSGVSLLIPAGA 558

Qy        526 IPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPD 585
Db        559 IPQGRVYEMYVTVHRKENMRPPMEDSQTLLTPVVSCGPPGALLTRPVILTLHHCADPSTE 618

Qy        586 SWSLRLKKQSCEGSWEQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA 645
Db        619 DWKIQLKNQAVQGQWE-DVVVVGEENFTTPCYIQLDAEACHILTENLSTYALVGQSTTKA 677

Qy        646 AAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDS 705
Db        678 AAKRLKLAIFGPLCCSSLEYSIRVYCLDDTQDALKEVLQLERQMGGQLLEEPKALHFKGS 737

Qy        706 YHNLRLSIHDVPSSLWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKL 765
Db        738 IHNLRLSIHDIAHSLWKSKLLAKYQEIPFYHIWSGSQRNLHCTFTLERLSLNTVELVCKL 797

Qy        766 WVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLD 825
Db        798 CVRQVEGEGQIFQLNCTVSEEPTGIDLPLLDPASTITTVTGPSAFSIPLPIRQKLCSSLD 857

Qy        826 PPCRRGADWRTLAQKLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLG 885
Db        858 APQTRGHDWRMLAHKLNLDRYLNYFATKSSPTGVILDLWEAQNFPDGNLSMLAAVLEEMG 917

Qy        886 QPDAGLFTVSEAE 898
Db        918 RHETVVSLAAEGQ 930

Search completed: October  6, 2004, 18:22:28
Job time : 103 secs
```

```
                    GenCore version 5.1.6
         Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - protein search, using sw model

Run on:          October  6, 2004, 18:16:33 ; Search time 27 Seconds
                                          (without alignments)
                                          1718.954 Million cell updates/sec

Title:           US-09-970-944-2
Perfect score:   4787
Sequence:        1 MAVRPGLWPALLGIVLAAWL..........AVAGLGQPDAGLFTVSEAEC 899

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters:      389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Issued_Patents_AA:*
                1:  /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
                2:  /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
                3:  /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
                4:  /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
                5:  /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
                6:  /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.
```

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 4545.5 | 95.0 | 898 | 2 | US-08-808-982-5 | Sequence 5, Appli |
| 2 | 4545.5 | 95.0 | 898 | 3 | US-09-306-902A-5 | Sequence 5, Appli |
| 3 | 2789.5 | 58.3 | 557 | 2 | US-08-808-982-6 | Sequence 6, Appli |
| 4 | 2789.5 | 58.3 | 557 | 3 | US-09-306-902A-6 | Sequence 6, Appli |
| 5 | 2490 | 52.0 | 943 | 2 | US-08-808-982-7 | Sequence 7, Appli |
| 6 | 2490 | 52.0 | 943 | 3 | US-09-306-902A-7 | Sequence 7, Appli |
| 7 | 294 | 6.1 | 102 | 2 | US-08-808-982-8 | Sequence 8, Appli |
| 8 | 294 | 6.1 | 102 | 3 | US-09-306-902A-8 | Sequence 8, Appli |
| 9 | 273.5 | 5.7 | 1172 | 1 | US-08-313-288B-19 | Sequence 19, Appl |
| 10 | 249 | 5.2 | 441 | 3 | US-08-985-526-3 | Sequence 3, Appli |
| 11 | 246 | 5.1 | 239 | 5 | PCT-US93-01652-1 | Sequence 1, Appli |
| 12 | 246 | 5.1 | 1170 | 1 | US-08-313-288B-20 | Sequence 20, Appl |
| 13 | 235.5 | 4.9 | 218 | 3 | US-08-985-526-1 | Sequence 1, Appli |
| 14 | 234 | 4.9 | 1395 | 3 | US-09-540-245A-15 | Sequence 15, Appl |
| 15 | 216.5 | 4.5 | 1651 | 3 | US-09-540-245A-18 | Sequence 18, Appl |
| 16 | 208.5 | 4.4 | 469 | 1 | US-08-313-288B-15 | Sequence 15, Appl |
| 17 | 205.5 | 4.3 | 1381 | 3 | US-09-540-245A-16 | Sequence 16, Appl |
| 18 | 198.5 | 4.1 | 788 | 2 | US-08-918-914-4 | Sequence 4, Appli |
| 19 | 194.5 | 4.1 | 1069 | 4 | US-09-877-730-2 | Sequence 2, Appli |
| 20 | 194.5 | 4.1 | 1150 | 4 | US-09-877-730-8 | Sequence 8, Appli |
| 21 | 189 | 3.9 | 904 | 4 | US-09-877-730-6 | Sequence 6, Appli |
| 22 | 189 | 3.9 | 985 | 4 | US-09-877-730-10 | Sequence 10, Appl |
| 23 | 188 | 3.9 | 1297 | 3 | US-09-540-245A-17 | Sequence 17, Appl |
| 24 | 186 | 3.9 | 380 | 4 | US-09-877-730-4 | Sequence 4, Appli |
| 25 | 186 | 3.9 | 1266 | 4 | US-08-506-296B-4 | Sequence 4, Appli |
| 26 | 168 | 3.5 | 481 | 4 | US-09-130-491-8 | Sequence 8, Appli |
| 27 | 167.5 | 3.5 | 905 | 4 | US-09-369-364A-9 | Sequence 9, Appli |
| 28 | 166 | 3.5 | 837 | 4 | US-09-122-126B-2 | Sequence 2, Appli |
| 29 | 166 | 3.5 | 837 | 4 | US-09-634-286A-2 | Sequence 2, Appli |
| 30 | 160.5 | 3.4 | 949 | 4 | US-09-568-559-2 | Sequence 2, Appli |
| 31 | 160.5 | 3.4 | 967 | 4 | US-09-130-491-2 | Sequence 2, Appli |
| 32 | 159.5 | 3.3 | 727 | 4 | US-09-445-023A-1 | Sequence 1, Appli |
| 33 | 159.5 | 3.3 | 1224 | 4 | US-09-930-872-4 | Sequence 4, Appli |
| 34 | 158 | 3.3 | 608 | 4 | US-09-130-491-13 | Sequence 13, Appl |
| 35 | 156 | 3.3 | 997 | 4 | US-09-369-364A-7 | Sequence 7, Appli |
| 36 | 153.5 | 3.2 | 1059 | 4 | US-09-907-794A-290 | Sequence 290, App |
| 37 | 153.5 | 3.2 | 1059 | 4 | US-09-905-125A-290 | Sequence 290, App |
| 38 | 153.5 | 3.2 | 1059 | 4 | US-09-902-775A-290 | Sequence 290, App |
| 39 | 153.5 | 3.2 | 1119 | 4 | US-09-907-794A-294 | Sequence 294, App |
| 40 | 153.5 | 3.2 | 1119 | 4 | US-09-905-125A-294 | Sequence 294, App |
| 41 | 153.5 | 3.2 | 1119 | 4 | US-09-902-775A-294 | Sequence 294, App |
| 42 | 153 | 3.2 | 1081 | 4 | US-09-369-364A-17 | Sequence 17, Appl |
| 43 | 152.5 | 3.2 | 757 | 4 | US-09-963-791-24 | Sequence 24, Appl |
| 44 | 152.5 | 3.2 | 908 | 4 | US-09-963-791-2 | Sequence 2, Appli |
| 45 | 152.5 | 3.2 | 1911 | 1 | US-08-348-006B-5 | Sequence 5, Appli |

ALIGNMENTS

```
RESULT 1
US-08-808-982-5
; Sequence 5, Application US/08808982
; Patent No. 5939271
;  GENERAL INFORMATION:
;    APPLICANT:  Tessier-Lavigne, Marc
;    APPLICANT:  Leonardo, E. David
;    APPLICANT:  Hink, Lindsay
;    APPLICANT:  Masu, Masayuki
;    APPLICANT:  Kazuko, Keino-Masu
;    TITLE OF INVENTION:  Netrin Receptors
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SCIENCE & TECHNOLOGY LAW GROUP
;      STREET:  268 BUSH STREET, SUITE 3200
;      CITY:  SAN FRANCISCO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/808,982
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  OSMAN, RICHARD A
;      REGISTRATION NUMBER:  36,627
;      REFERENCE/DOCKET NUMBER:  UC96-217
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 343-4341
;      TELEFAX:  (415) 343-4342
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  898 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  peptide
US-08-808-982-5

  Query Match            95.0%;  Score 4545.5;  DB 2;  Length 898;
  Best Local Similarity  94.7%;  Pred. No. 0;
  Matches  852;  Conservative   19;  Mismatches   26;  Indels    3;  Gaps    3;

Qy          1 MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60
                |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db          1 MAVRPGLWPVLLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60

Qy         61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGEPTMEVRINVSRQQVEKVFGLE 120
              ||||||||||||||||||||||||||||||||||| ||| ||||||||||||||||||||
Db         61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDSSSGLPTMEVRINVSRQQVEKVFGLE 120

Qy        121 EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180
              ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        121 EYWCQCVAWSSSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180

Qy        181 VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        181 VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSTSAAVIVY 240

Qy        241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDG 300
              ||||||||||||||||||||||||||||||||||||||||||||||  :|  : |  |||
Db        241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNV-QKTACATLCPVDG 299

Qy        301 SWSPWSKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCVHSASGPEDV 360
              ||| |||||||||||||||||||||||||||||||:| ||||||||||||:|:|| ||||
Db        300 SWSSWSKWSACGLDCTHWRSRECSDPAPRNGGEECRGADLDTRNCTSDLCLHTASCPEDV 359

Qy        361 ALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHL 420
              |||:||:|||||| |||| | |:|||||||||||||||||||||||||||||||||||||
Db        360 ALYIGLVAVAVCLFLLLLALGLIYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHL 419

Qy        421 LTIQPDLS-TTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEEFV 479
              |||||||| |||||||||| |||||||||||:||||||||| |||||||||||||||:||
Db        420 LTIQPDLSTTTTTYQGSLCSRQDGPSPKFQLSNGHLLSPLGSGRHTLHHSSPTSEAEDFV 479

Qy        480 SRLSTQNYFRSLPRGTSNMTYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLH 539
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        480 SRLSTQNYFRSLPRGTSNMAYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLH 539

Qy        540 KPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGS 599
              ||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||
Db        540 KPEDVRLPLAGCQTLLSPVVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGS 599

Qy        600 WEQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAAKRLKLLLFAPVA 659
              || ||||||||:||||||||||| ||||||||||||||||||||||| |||:||||||||
Db        600 WE-DVLHLGEESPSHLYYCQLEAGACYVFTEQLGRFALVGEALSVAATKRLRLLLFAPVA 658

Qy        660 CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSS 719
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        659 CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSS 718

Qy        720 LWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSI 779
              |||||||||||||||||||||||:||||||||||:: ||||||||:||||||||||||:|
Db        719 LWKSKLLVSYQEIPFYHIWNGTQQYLHCTFTLERINASTSDLACKVWVWQVEGDGQSFNI 778

Qy        780 NFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQ 839
              ||||||||||||||||||| ||||||||||||||||||||||:|||||| ||||||||||
Db        779 NFNITKDTRFAELLALESEGGVPALVGPSAFKIPFLIRQKIIASLDPPCSRGADWRTLAQ 838

Qy        840 KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 899
              ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        839 KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLGQLAAAVAGLGQPDAGLFTVSEAEC 898
```

```
RESULT 2
US-09-306-902A-5
; Sequence 5, Application US/09306902A
; Patent No. 6277585
;  GENERAL INFORMATION:
;       APPLICANT: Tessier-Lavigne, Marc
;                  Leonardo, E. David
;                  Hink, Lindsay
;                  Masu, Masayuki
;                  Kazuko, Keino-Masu
;       TITLE OF INVENTION: Netrin Receptors
;       NUMBER OF SEQUENCES: 9
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
;            STREET: 268 BUSH STREET, SUITE 3200
;            CITY: SAN FRANCISCO
;            STATE: CALIFORNIA
;            COUNTRY: USA
;            ZIP: 94104
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/306,902A
;            FILING DATE: 07-May-1999
;            CLASSIFICATION: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: OSMAN, RICHARD A
;            REGISTRATION NUMBER: 36,627
;            REFERENCE/DOCKET NUMBER: UC96-217
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (415) 343-4341
;            TELEFAX: (415) 343-4342
;  INFORMATION FOR SEQ ID NO: 5:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 898 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: not relevant
;            TOPOLOGY: not relevant
;       MOLECULE TYPE: peptide
;       SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-306-902A-5

  Query Match             95.0%;  Score 4545.5;  DB 3;  Length 898;
  Best Local Similarity   94.7%;  Pred. No. 0;
  Matches  852;  Conservative   19;  Mismatches   26;  Indels    3;  Gaps    3;

Qy          1 MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60
              ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAVRPGLWPVLLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60

Qy         61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGEPTMEVRINVSRQQVEKVFGLE 120
              ||||||||||||||||||||||||||||||||||| ||| ||||||||||||||||||||
Db         61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDSSSGLPTMEVRINVSRQQVEKVFGLE 120

Qy        121 EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180
              ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        121 EYWCQCVAWSSSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180

Qy        181 VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        181 VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSTSAAVIVY 240

Qy        241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDG 300
              ||||||||||||||||||||||||||||||||||||||||||||||  :|  : |  |||
Db        241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNV-QKTACATLCPVDG 299

Qy        301 SWSPWSKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCVHSASGPEDV 360
              ||| |||||||||||||||||||||||||||||||:| ||||||||||||:|:|| ||||
Db        300 SWSSWSKWSACGLDCTHWRSRECSDPAPRNGGEECRGADLDTRNCTSDLCLHTASCPEDV 359

Qy        361 ALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHL 420
              |||:||:|||||| |||| | |:|||||||||||||||||||||||||||||||||||||
Db        360 ALYIGLVAVAVCLFLLLLALGLIYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHL 419

Qy        421 LTIQPDLS-TTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEEFV 479
              |||||||| |||||||||| |||||||||||:||||||||| |||||||||||||||:||
Db        420 LTIQPDLSTTTTTYQGSLCSRQDGPSPKFQLSNGHLLSPLGSGRHTLHHSSPTSEAEDFV 479

Qy        480 SRLSTQNYFRSLPRGTSNMTYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLH 539
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        480 SRLSTQNYFRSLPRGTSNMAYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLH 539

Qy        540 KPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGS 599
              ||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||
```

```
Db        540 KPEDVRLPLAGCQTLLSPVVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGS 599

Qy        600 WEQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAAKRLKLLLFAPVA 659
              || ||||||||:||||||||||| ||||||||||||||||||||||| ||:||||||||
Db        600 WE-DVLHLGEESPSHLYYCQLEAGACYVFTEQLGRFALVGEALSVAATKRLRLLLFAPVA 658

Qy        660 CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSS 719
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        659 CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSS 718

Qy        720 LWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSI 779
              |||||||||||||||||||||||:||||||||||::: |||||||:||||||||||||:|
Db        719 LWKSKLLVSYQEIPFYHIWNGTQQYLHCTFTLERINASTSDLACKVWVWQVEGDGQSFNI 778

Qy        780 NFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQ 839
              ||||||||||||||||||| |||||||||||||||||||||||:||||| ||||||||||
Db        779 NFNITKDTRFAELLALESEGGVPALVGPSAFKIPFLIRQKIIASLDPPCSRGADWRTLAQ 838

Qy        840 KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 899
              |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db        839 KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLGQLAAAVAGLGQPDAGLFTVSEAEC 898
```

RESULT 3
US-08-808-982-6
; Sequence 6, Application US/08808982
; Patent No. 5939271
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; APPLICANT: Leonardo, E. David
; APPLICANT: Hink, Lindsay
; APPLICANT: Masu, Masayuki
; APPLICANT: Kazuko, Keino-Masu
; TITLE OF INVENTION: Netrin Receptors
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/808,982
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 557 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
US-08-808-982-6

```
  Query Match            58.3%;  Score 2789.5;  DB 2;  Length 557;
  Best Local Similarity  96.4%;  Pred. No. 2.7e-256;
  Matches  538;  Conservative    2;  Mismatches   15;  Indels    3;  Gaps    3;

Qy        344 NCTSDLCVHSASGPEDVALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTS 403
              |||||| ||:||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 NCTSDLXVHTASGPEDVALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTS 60

Qy        404 GFQPVSIKPSKADNPHLLTIQPDLS-TTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGG 462
              ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db         61 GFQPVSIKPSKADNPHLLTIQPDLSTTTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGG 120

Qy        463 RHTLHHSSPTSEAEEFVSRLSTQNYFRSLPRGTSNMTYGTFNFLGGRLMIPNTGISLLIP 522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RHTLHHSSPTSEAEEFVSRLSTQNYFRSLPRGTSNMTYGTFNFLGGRLMIPNTGISLLIP 180

Qy        523 PDAIPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEP 582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PDAIPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEP 240

Qy        583 SPDSWSLRLKKQSCEGSWEQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEAL 642
              ||||||| ||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        241 SPDSWSLALKKQSCEGSWE-DVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEAL 299

Qy        643 SVAAAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHF 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db        300 SVAAAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHL 359

Qy        703 KDSYHNLRLSIHDVPSSLWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLA 762
               |||||| || |||||||||||||||||||||||||||||||||||||||||||||||||
Db        360 XDSYHNLXLSXHDVPSSLWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLA 419

Qy        763 CKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIIS 822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        420 CKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIIS 479

Qy        823 SLDPPCRRGADWRTLAQKLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVA 882
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 SLDPPCRRGADWRTLAQKLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVA 539

Qy        883 GLGQPDAGLFT-VSEAEC 899
              |       |:   |||||
Db        540 GTXPAGRWLLSQCSEAEC 557
```

RESULT 4
US-09-306-902A-6
; Sequence 6, Application US/09306902A
; Patent No. 6277585
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; Leonardo, E. David
; Hink, Lindsay
; Masu, Masayuki
; Kazuko, Keino-Masu
; TITLE OF INVENTION: Netrin Receptors
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/306,902A
; FILING DATE: 07-May-1999
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341

```
;            TELEFAX: (415) 343-4342
;   INFORMATION FOR SEQ ID NO: 6:
;         SEQUENCE CHARACTERISTICS:
;            LENGTH: 557 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: not relevant
;            TOPOLOGY: not relevant
;         MOLECULE TYPE: peptide
;         SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-306-902A-6

  Query Match            58.3%;  Score 2789.5;  DB 3;  Length 557;
  Best Local Similarity  96.4%;  Pred. No. 2.7e-256;
  Matches  538;  Conservative    2;  Mismatches   15;  Indels    3;  Gaps    3;

Qy        344 NCTSDLCVHSASGPEDVALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTS 403
Db          1 NCTSDLXVHTASGPEDVALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTS 60

Qy        404 GFQPVSIKPSKADNPHLLTIQPDLS-TTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGG 462
Db         61 GFQPVSIKPSKADNPHLLTIQPDLSTTTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGG 120

Qy        463 RHTLHHSSPTSEAEEFVSRLSTQNYFRSLPRGTSNMTYGTFNFLGGRLMIPNTGISLLIP 522
Db        121 RHTLHHSSPTSEAEEFVSRLSTQNYFRSLPRGTSNMTYGTFNFLGGRLMIPNTGISLLIP 180

Qy        523 PDAIPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEP 582
Db        181 PDAIPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEP 240

Qy        583 SPDSWSLRLKKQSCEGSWEQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEAL 642
Db        241 SPDSWSLALKKQSCEGSWE-DVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEAL 299

Qy        643 SVAAAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHF 702
Db        300 SVAAAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHL 359

Qy        703 KDSYHNLRLSIHDVPSSLWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLA 762
Db        360 XDSYHNLXLSXHDVPSSLWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLA 419

Qy        763 CKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIIS 822
Db        420 CKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIIS 479

Qy        823 SLDPPCRRGADWRTLAQKLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVA 882
Db        480 SLDPPCRRGADWRTLAQKLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVA 539

Qy        883 GLGQPDAGLFT-VSEAEC 899
Db        540 GTXPAGRWLLSQCSEAEC 557

RESULT 5
US-08-808-982-7
; Sequence 7, Application US/08808982
; Patent No. 5939271
;  GENERAL INFORMATION:
;    APPLICANT:  Tessier-Lavigne, Marc
;    APPLICANT:  Leonardo, E. David
;    APPLICANT:  Hink, Lindsay
;    APPLICANT:  Masu, Masayuki
;    APPLICANT:  Kazuko, Keino-Masu
;    TITLE OF INVENTION:  Netrin Receptors
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SCIENCE & TECHNOLOGY LAW GROUP
;      STREET:  268 BUSH STREET, SUITE 3200
;      CITY:  SAN FRANCISCO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/808,982
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  OSMAN, RICHARD A
;      REGISTRATION NUMBER:  36,627
;      REFERENCE/DOCKET NUMBER:  UC96-217
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 343-4341
;      TELEFAX:  (415) 343-4342
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  943 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  peptide
US-08-808-982-7

  Query Match            52.0%;  Score 2490;  DB 2;  Length 943;
  Best Local Similarity  52.3%;  Pred. No. 2e-227;
  Matches  496;  Conservative  144;  Mismatches  226;  Indels   82;  Gaps   19;

Qy          9 PALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAV 68
Db         21 PSLAGI-------DSGAQ---GLPDSFPSAPAEQLPHFLLEPEDAYIVKNKPVELHCRAF 70

Qy         69 PATQIFFKCNGEWVRQVDHVIERSTDGSSGEPTMEVRINVSRQQVEKVFGLEEYWCQCVA 128
Db         71 PATQIYFKCNGEWVSQKGHVTQESLDEATGLRIREVQIEVSRQQVEELFGLEDYWCQCVA 130

Qy        129 WSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAEVEWLRNED 188
Db        131 WSSSGTTKSRRAYIRIAYLRKNFDQEPLAKEVPLDHEVLLQCRPPEGVPVAEVEWLKNED 190

Qy        189 LVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNGGWSTW 248
Db        191 VIDPAQDTNFLLTIDHNLIIRQARLSDTANYTCVAKNIVAKRRSTTATVIVYVNGGWSSW 250

Qy        249 TEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDGSWSPWSKW 308
Db        251 AEWSPCSNRCGRGWQKRTRTCTNPAPLNGGAFCEGQ-ACQKTACTTVCPVDGAWTEWSKW 309

Qy        309 SACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCV---HSASGPE------- 358
Db        310 SACSTECAHWRSRECMAPPPQNGGRDCSGTLLDSKNCTDGLCVLNQRTLNDPKSRPLEPS 369

Qy        359 -DVALYVGL-IAVAVCLVLLLLVLILVYCRKKEGLDSDVADSS-ILTSGFQPVSIKPSKA 415
Db        370 GDVALYAGLVVAVFVVLAVLMAVGVIVYRRNCRDFDTDITDSSAALTGGFHPVNFKTARP 429

Qy        416 DNPHLL--TIQPDLSTTT-TYQGSLCPRQDGPSPKFQLTNGHLLSPL------------- 459
Db        430 SNPQLLHPSAPPDLTASAGIYRGPVYALQDS-ADKIPMTNSPLLDPLPSLKIKVYDSSTI 488

Qy        460 --GGG----------------------RHTLHHSSPTSEAEEFVSRLSTQNYFRSLPRG 494
Db        489 GSGAGLADGADLLGVLPPGTYPGDFSRDTHFLHLRS---------ASLGSQ-HLLGLPRD 538

Qy        495 TSNMTYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLHKPEDVRLPLA-GCQT 553
Db        539 PSSSVSGTFGCLGGRLTIPGTGVSLLVPNGAIPQGKFYDLYLRINKTEST-LPLSEGSQT 597

Qy        554 LLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGSWEQDVLHLGEEAPS 613
```

```
Db    598 VLSPSVTCGPTGLLLCRPVVLTVPHCAEVIAGDWIFQLKTQAHQGHWE-EVVTLDEETLN 656

Qy    614 HLYYCQLEASACYVFTEQLGRFALVGEALSVAAAKRLKLLLFAPVACTSLEYNIRVYCLH 673
Db    657 TPCYCQLEAKSCHILLDQLGTYVFTGESYSRSAVKRLQLAIFAPALCTSLEYSLRVYCLE 716

Qy    674 DTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWKSKLLVSYQEIP 733
Db    717 DTPAALKEVLELERTLGGYLVEEPKTLLFKDSYHNLRLSLHDIPHAHWRSKLLAKYQEIP 776

Qy    734 FYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELL 793
Db    777 FYHVWNGSQKALHCTFTLERHSLASTEFTCKVCVRQVEGEGQIFQLHTTLA-ETPAGSLD 835

Qy    794 ALESEAGVPAL--VGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLDSHLSFFA 851
Db    836 ALCSAPGNAATTQLGPYAFKIPLSIRQKICNSLDAPNSRGNDWRLLAQKLSMDRYLNYFA 895

Qy    852 SKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 899
Db    896 TKASPTGVILDLWEARQQDDGDLNSLASALEEMGKSEMLVAMTTDGDC 943
```

RESULT 6
US-09-306-902A-7
; Sequence 7, Application US/09306902A
; Patent No. 6277585
;   GENERAL INFORMATION:
;        APPLICANT: Tessier-Lavigne, Marc
;                   Leonardo, E. David
;                   Hink, Lindsay
;                   Masu, Masayuki
;                   Kazuko, Keino-Masu
;        TITLE OF INVENTION: Netrin Receptors
;        NUMBER OF SEQUENCES: 9
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
;             STREET: 268 BUSH STREET, SUITE 3200
;             CITY: SAN FRANCISCO
;             STATE: CALIFORNIA
;             COUNTRY: USA
;             ZIP: 94104
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/306,902A
;             FILING DATE: 07-May-1999
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: OSMAN, RICHARD A
;             REGISTRATION NUMBER: 36,627
;             REFERENCE/DOCKET NUMBER: UC96-217
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 343-4341
;             TELEFAX: (415) 343-4342
;   INFORMATION FOR SEQ ID NO: 7:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 943 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: not relevant
;             TOPOLOGY: not relevant
;        MOLECULE TYPE: peptide
;        SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-306-902A-7

```
  Query Match             52.0%;  Score 2490;  DB 3;  Length 943;
  Best Local Similarity   52.3%;  Pred. No. 2e-227;
  Matches  496;  Conservative  144;  Mismatches  226;  Indels   82;  Gaps   19;

Qy      9 PALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAV 68
Db     21 PSLAGI-------DSGAQ---GLPDSFPSAPAEQLPHFLLEPEDAYIVKNKPVELHCRAF 70

Qy     69 PATQIFFKCNGEWVRQVDHVIERSTDGSSGEPTMEVRINVSRQQVEKVFGLEEYWCQCVA 128
Db     71 PATQIYFKCNGEWVSQKGHVTQESLDEATGLRIREVQIEVSRQQVEELFGLEDYWCQCVA 130

Qy    129 WSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAEVEWLRNED 188
Db    131 WSSSGTTKSRRAYIRIAYLRKNFDQEPLAKEVPLDHEVLLQCRPPEGVPVAEVEWLKNED 190

Qy    189 LVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNGGWSTW 248
Db    191 VIDPAQDTNFLLTIDHNLIIRQARLSDTANYTCVAKNIVAKRRSTTATVIVYVNGGWSSW 250

Qy    249 TEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDGSWSPWSKW 308
Db    251 AEWSPCSNRCGRGWQKRTRTCTNPAPLNGGAFCEGQ-ACQKTACTTVCPVDGAWTEWSKW 309

Qy    309 SACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCV---HSASGPE------- 358
Db    310 SACSTECAHWRSRECMAPPPQNGGRDCSGTLLDSKNCTDGLCVLNQRTLNDPKSRPLEPS 369

Qy    359 -DVALYVGL-IAVAVCLVLLLLVLILVYCRKKEGLDSDVADSS-ILTSGFQPVSIKPSKA 415
Db    370 GDVALYAGLVVAVFVVLAVLMAVGVIVYRRNCRDFDTDITDSSAALTGGFHPVNFKTARP 429

Qy    416 DNPHLL--TIQPDLSTTT-TYQGSLCPRQDGPSPKFQLTNGHLLSPL------------ 459
Db    430 SNPQLLHPSAPPDLTASAGIYRGPVYALQDS-ADKIPMTNSPLLDPLPSLKIKVYDSSTI 488

Qy    460 --GGG---------------------RHTLHHSSPTSEAEEFVSRLSTQNYFRSLPRG 494
Db    489 GSGAGLADGADLLGVLPPGTYPGDFSRDTHFLHLRS---------ASLGSQ-HLLGLPRD 538

Qy    495 TSNMTYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLHKPEDVRLPLA-GCQT 553
Db    539 PSSSVSGTFGCLGGRLTIPGTGVSLLVPNGAIPQGKFYDLYLRINKTEST-LPLSEGSQT 597

Qy    554 LLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGSWEQDVLHLGEEAPS 613
Db    598 VLSPSVTCGPTGLLLCRPVVLTVPHCAEVIAGDWIFQLKTQAHQGHWE-EVVTLDEETLN 656

Qy    614 HLYYCQLEASACYVFTEQLGRFALVGEALSVAAAKRLKLLLFAPVACTSLEYNIRVYCLH 673
Db    657 TPCYCQLEAKSCHILLDQLGTYVFTGESYSRSAVKRLQLAIFAPALCTSLEYSLRVYCLE 716

Qy    674 DTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWKSKLLVSYQEIP 733
Db    717 DTPAALKEVLELERTLGGYLVEEPKTLLFKDSYHNLRLSLHDIPHAHWRSKLLAKYQEIP 776

Qy    734 FYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELL 793
Db    777 FYHVWNGSQKALHCTFTLERHSLASTEFTCKVCVRQVEGEGQIFQLHTTLA-ETPAGSLD 835

Qy    794 ALESEAGVPAL--VGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLDSHLSFFA 851
Db    836 ALCSAPGNAATTQLGPYAFKIPLSIRQKICNSLDAPNSRGNDWRLLAQKLSMDRYLNYFA 895

Qy    852 SKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 899
Db    896 TKASPTGVILDLWEARQQDDGDLNSLASALEEMGKSEMLVAMTTDGDC 943
```

RESULT 7
US-08-808-982-8
; Sequence 8, Application US/08808982
; Patent No. 5939271
;   GENERAL INFORMATION:
;     APPLICANT:  Tessier-Lavigne, Marc
;     APPLICANT:  Leonardo, E. David
;     APPLICANT:  Hink, Lindsay
;     APPLICANT:  Masu, Masayuki

```
;    APPLICANT:  Kazuko, Keino-Masu
;    TITLE OF INVENTION:  Netrin Receptors
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SCIENCE & TECHNOLOGY LAW GROUP
;      STREET:  268 BUSH STREET, SUITE 3200
;      CITY:  SAN FRANCISCO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/808,982
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  OSMAN, RICHARD A
;      REGISTRATION NUMBER:  36,627
;      REFERENCE/DOCKET NUMBER:  UC96-217
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 343-4341
;      TELEFAX:  (415) 343-4342
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  102 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  peptide
US-08-808-982-8

  Query Match             6.1%;  Score 294;  DB 2;  Length 102;
  Best Local Similarity   56.4%;  Pred. No. 4.1e-20;
  Matches   57;  Conservative   16;  Mismatches   28;  Indels    0;  Gaps    0;

Qy        609 EEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAAKRLKLLLFAPVACTSLEYNIR 668
              ||   :     | |||  ||  :   :|||  :      ||:  |  :|  |||:|  :|||   ||||||::|
Db          2 EETLNTPCYXQLEPRACXILLDQLGTYVFTGESYSRSAVKRLQLAVFAPALCTSLEYSLR 61

Qy        669 VYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNL 709
              ||||  ||    |||||::||:  |||  |::||:  |  ||||||||
Db         62 VYCLEDTPVALKEVLELERTLGGYLVEEPKPLMFKDSYHNL 102


RESULT 8
US-09-306-902A-8
; Sequence 8, Application US/09306902A
; Patent No. 6277585
;   GENERAL INFORMATION:
;        APPLICANT: Tessier-Lavigne, Marc
;                   Leonardo, E. David
;                   Hink, Lindsay
;                   Masu, Masayuki
;                   Kazuko, Keino-Masu
;        TITLE OF INVENTION: Netrin Receptors
;        NUMBER OF SEQUENCES: 9
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
;             STREET: 268 BUSH STREET, SUITE 3200
;             CITY: SAN FRANCISCO
;             STATE: CALIFORNIA
;             COUNTRY: USA
;             ZIP: 94104
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/306,902A
;             FILING DATE: 07-May-1999
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: OSMAN, RICHARD A
;             REGISTRATION NUMBER: 36,627
;             REFERENCE/DOCKET NUMBER: UC96-217
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 343-4341
;             TELEFAX: (415) 343-4342
;   INFORMATION FOR SEQ ID NO: 8:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 102 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: not relevant
;             TOPOLOGY: not relevant
;        MOLECULE TYPE: peptide
;        SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-306-902A-8

  Query Match             6.1%;  Score 294;  DB 3;  Length 102;
  Best Local Similarity   56.4%;  Pred. No. 4.1e-20;
  Matches   57;  Conservative   16;  Mismatches   28;  Indels    0;  Gaps    0;

Qy        609 EEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAAKRLKLLLFAPVACTSLEYNIR 668
              ||   :     | |||  ||  :   :|||  :      ||:  |  :|  |||:|  :|||   ||||||::|
Db          2 EETLNTPCYXQLEPRACXILLDQLGTYVFTGESYSRSAVKRLQLAVFAPALCTSLEYSLR 61

Qy        669 VYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNL 709
              ||||  ||    |||||::||:  |||  |::||:  |  ||||||||
Db         62 VYCLEDTPVALKEVLELERTLGGYLVEEPKPLMFKDSYHNL 102


RESULT 9
US-08-313-288B-19
; Sequence 19, Application US/08313288B
; Patent No. 5750502
;  GENERAL INFORMATION:
;    APPLICANT:  Jessell, Thomas M. and Avihu Klar
;    TITLE OF INVENTION:  CLONING, EXPRESSION AND USES OF A
;    TITLE OF INVENTION:  NOVEL SECRETED PROTEIN, F-SPONDIN
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cooper & Dunham LLP
;      STREET:  1185 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/313,288B
;      FILING DATE:  January 5, 1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  White, John P.
;      REGISTRATION NUMBER:  28,678
;      REFERENCE/DOCKET NUMBER:  40028-A-PCT-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 278-0400
;      TELEFAX:  (212) 391-0526
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  19:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1172 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
```

```
US-08-313-288B-19

  Query Match             5.7%;  Score 273.5;  DB 1;  Length 1172;
  Best Local Similarity   28.9%;  Pred. No. 2.1e-16;
  Matches   74;  Conservative   27;  Mismatches  110;  Indels   45;  Gaps    8;

Qy        209 RQARLADTANYTCVAKNIVARRRSASAA-VIVYVNGGWSTWTEWSVCSASCGRGWQKRSR 267
Db        403 QRGRSCDVTSNTCLGPSIQTRACSLSKCDTRIRQDGGWSHWSPWSSCSVTCGVGNITRIR 462

Qy        268 SCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDGSWSPWSKWSACGLDCT---HWRSRECS 324
Db        463 LCNSPVPQMGGKNCKGSGRETKACQGAPCPIDGRWSPWSPWSACTVTCAGGIRERTRVCN 522

Qy        325 DPAPRNGGEECQGTDLDTRNCTSDLCVHSASGPEDVALYVGLIAVAVCLVLLLLVLILVY 384
Db        523 SPEPQYGGKACVGDVQERQMCNKRSC------PVDGCLSNPCFPGAQC------------ 564

Qy        385 CRKKEGLDSDVADSSILTSGFQPVSI--KPSKADNPHLLTIQPDLSTTT--------TYQ 434
Db        565 --------SSFPDGS-WSCGFCPVGFLGNGTHCEDLDECALVPDICFSTSKVPRCVNTQP 615

Qy        435 GSLC----PRQDGPSP 446
Db        616 GFHCLPCPPRYRGNQP 631

RESULT 10
US-08-985-526-3
; Sequence 3, Application US/08985526
; Patent No. 6080728
;  GENERAL INFORMATION:
;    APPLICANT:  Mixson, James A
;    TITLE OF INVENTION:  CARRIER:DNA COMPLEXES CONTAINING DNA
;    TITLE OF INVENTION:  ENCODING ANTI-ANGIOGENIC PEPTIDES AND THEIR USE IN GENE
;    TITLE OF INVENTION:  THERAPY
;    NUMBER OF SEQUENCES:  43
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Connolly, Bove, Lodge, & Hutz
;      STREET:  1220 Market Street, P.O. Box 2207
;      CITY:  Wilmington
;      STATE:  Delaware
;      COUNTRY:  U.S.A.
;      ZIP:  19899
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/985,526
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/608,845
;      FILING DATE:  16-JUL-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McMorrow Jr., Robert G
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (302) 658-9141
;      TELEFAX:  (302) 658-5613
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  441 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-985-526-3

  Query Match             5.2%;  Score 249;  DB 3;  Length 441;
  Best Local Similarity   25.9%;  Pred. No. 9e-15;
  Matches   87;  Conservative   35;  Mismatches  114;  Indels  100;  Gaps   15;

Qy         75 FKCNGEW----------VRQVDHVIERSTDGSSGEPTM-------EVRINVSRQQ----V 113
Db        132 FKQDGGWSHWSPWSSCSVTCGDGVITRITLCNSPSPQMNGKPCEGEARETKACKKDACPI 191

Qy        114 EKVFGLEEYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPP 173
Db        192 NGGWGPWSPWDICSVTCGGGVQKRSRLCV---DSRMTEENKELANELR---------RPP 239

Qy        174 ----EGIPPAEVEWLRNED-LVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVA 228
Db        240 LCYHNG-----VQYRNNEEWTVDSCTE-------------------------CHCQNSVT 269

Qy        229 RRRSASAAVIVYVNG-------------------GWSTWTEWSVCSASCGRGWQKRSRSC 269
Db        270 ICKKVSCPIMPCSNATVPDGECCPRCWPSDSADDGWSPWSEWTSCSTSCGNGIQQRGRSC 329

Qy        270 TNPAPLNGGAFCEGQNVHDRTV----SSLLVSVDGSWSPWSKWSACGLDC---THWRSRE 322
Db        330 DS---LNNR--CEGSSVQTRTCHIQECDKRFKQDGGWSHWSPWSSCSVTCGDGVITRITL 384

Qy        323 CSDPAPRNGGEECQGTDLDTRNCTSDLC-VHSASGP 357
Db        385 CNSPSPQMNGKPCEGEARETKACKKDACPINGGWGP 420

RESULT 11
PCT-US93-01652-1
; Sequence 1, Application PC/TUS9301652
;  GENERAL INFORMATION:
;    APPLICANT:  Bouck, Noel P.
;    APPLICANT:  Polverini, Peter J.
;    APPLICANT:  Good, Deborah J.
;    APPLICANT:  Frazier, William A.
;    TITLE OF INVENTION:  Method and Composition for
;    TITLE OF INVENTION:  Inhibiting Angiogenesis
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Tilton, Fallon, Lungmus & Chestnut
;      STREET:  100 South Wacker Drive, Suite 960
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606-4002
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/01652
;      FILING DATE:  19930222
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/841,656
;      FILING DATE:  24-FEB-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/464,369
;      FILING DATE:  12-JAN-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fentress, Susan B.
;      REGISTRATION NUMBER:  31,327
;      REFERENCE/DOCKET NUMBER:  92005-PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312)-456-8000
;      TELEFAX:  (312)-456-7776
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  239 amino acids
;      TYPE:  AMINO ACID
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
PCT-US93-01652-1
```

```
  Query Match             5.1%;  Score 246;  DB 5;  Length 239;
  Best Local Similarity   31.0%;  Pred. No. 6.2e-15;
  Matches   48;  Conservative   23;  Mismatches   66;  Indels   18;  Gaps    3;

Qy        207 VVRQARLADTANYTCVAKNIVAR--------RRSASAAVIVYVNGGWSTWTEWSVCSASC 258
              : ::  |  |: |  |   ::  |         :|        :|||  |: || || :|
Db         88 IQQRGRSCDSLNNRCEGSSVQTRTCHIQECDKRFKQ-------DGGWSHWSPWSSCSVTC 140

Qy        259 GRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDGSWSPWSKWSACGLDC--- 315
              | |    | | | :|:|   |  |||:      :          ::|  | |||  |  : |
Db        141 GDGVITRIRLCNSPSPQMNGLPCEGEARETKACKKDACPINGGWGPWSPWDICSVTCGGG 200

Qy        316 THWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLC 350
                  ||| |::|||: || :|  |    :  :  |     |
Db        201 VQKRSRLCNNPAPQFGGLDCVGDVTENQICNKQDC 235
```

RESULT 12

```
US-08-313-288B-20
; Sequence 20, Application US/08313288B
; Patent No. 5750502
;  GENERAL INFORMATION:
;    APPLICANT:  Jessell, Thomas M. and Avihu Klar
;    TITLE OF INVENTION:  CLONING, EXPRESSION AND USES OF A
;    TITLE OF INVENTION:  NOVEL SECRETED PROTEIN, F-SPONDIN
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cooper & Dunham LLP
;      STREET:  1185 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/313,288B
;      FILING DATE:  January 5, 1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  White, John P.
;      REGISTRATION NUMBER:  28,678
;      REFERENCE/DOCKET NUMBER:  40028-A-PCT-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 278-0400
;      TELEFAX:  (212) 391-0526
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1170 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-313-288B-20

  Query Match             5.1%;  Score 246;  DB 1;  Length 1170;
  Best Local Similarity   30.3%;  Pred. No. 8.8e-14;
  Matches   47;  Conservative   24;  Mismatches   66;  Indels   18;  Gaps    3;

Qy        207 VVRQARLADTANYTCVAKNIVAR--------RRSASAAVIVYVNGGWSTWTEWSVCSASC 258
              : ::  |  |: |  |   ::  |         :|        :|||  |: || || :|
Db        399 IQQRGRSCDSLNNRCEGSSVQTRTCHIQECDKRFKQ-------DGGWSHWSPWSSCSVTC 451

Qy        259 GRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDGSWSPWSKWSACGLDC--- 315
              | |    | | | :|:|   |  |||:      :          ::|  | |||  |  : |
Db        452 GDGVITRIRLCNSPSPQMNGKPCEGEARETKACKKDACPINGGWGPWSPWDICSVTCGGG 511

Qy        316 THWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLC 350
                  ||| |::|  |: ||::| |    :  :  |     |
Db        512 VQKRSRLCNNPTPQFGGKDCVGDVTENQICNKQDC 546
```

RESULT 13

```
US-08-985-526-1
; Sequence 1, Application US/08985526
; Patent No. 6080728
;  GENERAL INFORMATION:
;    APPLICANT:  Mixson, James A
;    TITLE OF INVENTION:  CARRIER:DNA COMPLEXES CONTAINING DNA
;    TITLE OF INVENTION:  ENCODING ANTI-ANGIOGENIC PEPTIDES AND THEIR USE IN GENE
;    TITLE OF INVENTION:  THERAPY
;    NUMBER OF SEQUENCES:  43
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Connolly, Bove, Lodge, & Hutz
;      STREET:  1220 Market Street, P.O. Box 2207
;      CITY:  Wilmington
;      STATE:  Delaware
;      COUNTRY:  U.S.A.
;      ZIP:  19899
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/985,526
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/608,845
;      FILING DATE:  16-JUL-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McMorrow Jr., Robert G
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (302) 658-9141
;      TELEFAX:  (302) 658-5613
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  218 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-985-526-1

  Query Match             4.9%;  Score 235.5;  DB 3;  Length 218;
  Best Local Similarity   38.5%;  Pred. No. 5.3e-14;
  Matches   47;  Conservative   16;  Mismatches   46;  Indels   13;  Gaps    5;

Qy        244 GWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTV----SSLLVSVD 299
              ||| |:||: || ||| | |:| ||| :     ||     ||| :|  ||           |
Db         81 GWSPWSEWTSCSTSCGNGIQQRGRSCDS---LNNR--CEGSSVQTRTCHIQECDKRFKQD 135

Qy        300 GSWSPWSKWSACGLDC---THWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLC-VHSAS 355
              | || || ||:| : |      |        |:|:  |: |:| : :|: |  | | ::
Db        136 GGWSHWSPWSSCSVTCGDGVITRITNLCSPSPQMNGKPCEGREAETKACKKDACPINGGW 195

Qy        356 GP 357
              ||
Db        196 GP 197
```

RESULT 14

```
US-09-540-245A-15
; Sequence 15, Application US/09540245A
; Patent No. 6270984
; GENERAL INFORMATION:
;  APPLICANT:  Goodman, Corey
;  APPLICANT:  Kid, Thomas
;  APPLICANT:  Brose, Katja
;  APPLICANT:  Tessier-Lavigne, Marc
;  TITLE OF INVENTION:  Modulating Robo: Ligand Interactions
;  FILE REFERENCE:  B98-031-3
;  CURRENT APPLICATION NUMBER: US/09/540,245A
```

```
;  CURRENT FILING DATE:  2000-03-31
;  PRIOR APPLICATION NUMBER:  60/065,544
;  PRIOR FILING DATE:  1997-11-14
;  PRIOR APPLICATION NUMBER:  60/081,057
;  PRIOR FILING DATE:  1998-04-07
;  NUMBER OF SEQ ID NOS:  20
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 15
;   LENGTH: 1395
;   TYPE: PRT
;   ORGANISM: Drosophila melanogaster
US-09-540-245A-15

  Query Match             4.9%;  Score 234;  DB 3;  Length 1395;
  Best Local Similarity   20.5%;  Pred. No. 1.6e-12;
  Matches  188;  Conservative  103;  Mismatches  262;  Indels  364;  Gaps   46;

Qy          4 RPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGA-----NPDLLPHFLVEPEDVYIVKN 58
Db         28 RMWLLPAWLLLVLVA-------------SNGLPAVRGQYQSPRIIEH----PTDLVVKKN 70

Qy         59 KPVLLVCK--AVPATQIFFKCNGEWVRQVDHVIERSTDGSSGEPTMEVRINVSR------ 110
Db         71 EPATLNCKVEGKPEPTI------EWFK-------------DGEP---VSTNEKKSHRVQF 108

Qy        111 ----------QQVEKVFGLEEYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEV 160
Db        109 KDGALFFYRTMQGKKEQDGGEYW--CVAKNRVGQAVSRHASLQIAVLRDDFRVEPKDTRV 166

Qy        161 SLEQGIVLPCRPPEGIPPAEVEWLRN----EDLVDPSL--DPNVYITREHSLVVRQARLA 214
Db        167 AKGETALLECGPPKGIPEPTLIWIKDGVPLDDLKAMSFGASSRVRIVDGGNLLISNVEPI 226

Qy        215 DTANYTCVAKNIVARRRSASAAVIVYVN--------------GGWSTWTEWSVCSASCGR 260
Db        227 DEGNYKCIAQNLVGTRESSYAKLIVQVKPYFMKEPKDQVMLYGQTATF--------HCSV 278

Qy        261 G--------WQK----------------RSRSCTNPAPLNGGAF-CEGQN---------- 285
Db        279 GGDPPPKVLWKKEEGNIPVSRARILHDEKSLEISNITPTDEGTYVCEAHNNVGQISARAS 338

Qy        286 --VH-----DRTVSSLLVSVDG---------SWSP-----WSKWSACGL---DCTHWRSRE 322
Db        339 LIVHAPPNFTKRPSNKKVGLNGVVQLPCMASGNPPPSVFWTKEGVSTLMFPNSSHGRQYV 398

Qy        323 CSDPAPRNGGEECQGTDLDTRNCTSDLCVHSASGPEDVALYVGLIAVAVCLVLLLLVLIL 382
Db        399 AADGT-------LQITDV---------------RQEDEGYY-------VCSAF------- 422

Qy        383 VYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADN--PHLLTIQPDLSTTTTYQGSL--- 437
Db        423 -----------SVVDSSTVR-----VFLQVSSVDERPPPIIQIGP--ANQTLPKGSVATL 464

Qy        438 -CPRQDGPSPKFQ-LTNGHLLSPLGGGRHTLHHSSPTSEAEEFVSRLSTQNYFRSLPRGT 495
Db        465 PCRATGNPSPRIKWFHDGHAVQ--AGNRYSIIQGSSLRVDDLQLSDSGTYTCTASGERGE 522

Qy        496 SNMTYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLHKPEDVRLPLAGCQTLL 555
Db        523 TS----------------------------------WAATLTVEKPGSTSLHRAA----- 543

Qy        556 SPIVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLR-LKKQSCEGS--------------- 599
Db        544 DPSTYPAPPGT----PKVLNV------SRTSISLRWAKSQEKPGAVGPIIGYTVEYFSPD 593

Qy        600 ----WEQDVLHLGEEAPSHLYYCQLEASACYVF--------------------TEQLGR 634
Db        594 LQTGWIVAAHRVGD---TQVTISGLTPGTSYVFLVRAENTQGISVPSGLSNVIKTIEADF 650

Qy        635 FALVGEALSVA----AAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLG 690
Db        651 DAASANDLSAARTLLTGKSVELIDASAINASAVRLE---WMLHVSAD--------EKYVE 699

Qy        691 GQLIQEPRVLHFKDS------YHNLRL---------------------------SIHDV 716
Db        700 GLRI------HYKDASVPSAQYHSITVMDASAESFVVGNLKKYTKYEFFLTPFFETIEGQ 753

Qy        717 PSSLWKSKLLVSYQEIP 733
Db        754 PSN---SKTALTYEDVP 767

RESULT 15
US-09-540-245A-18
; Sequence 18, Application US/09540245A
; Patent No. 6270984
; GENERAL INFORMATION:
;  APPLICANT:  Goodman, Corey
;  APPLICANT:  Kid, Thomas
;  APPLICANT:  Brose, Katja
;  APPLICANT:  Tessier-Lavigne, Marc
;  TITLE OF INVENTION:  Modulating Robo: Ligand Interactions
;  FILE REFERENCE:  B98-031-3
;  CURRENT APPLICATION NUMBER: US/09/540,245A
;  CURRENT FILING DATE:  2000-03-31
;  PRIOR APPLICATION NUMBER:  60/065,544
;  PRIOR FILING DATE:  1997-11-14
;  PRIOR APPLICATION NUMBER:  60/081,057
;  PRIOR FILING DATE:  1998-04-07
;  NUMBER OF SEQ ID NOS:  20
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 18
;   LENGTH: 1651
;   TYPE: PRT
;   ORGANISM: human
US-09-540-245A-18

  Query Match             4.5%;  Score 216.5;  DB 3;  Length 1651;
  Best Local Similarity   20.7%;  Pred. No. 1e-10;
  Matches  213;  Conservative  139;  Mismatches  381;  Indels  295;  Gaps   54;

Qy         41 DLLPHFLVEPEDVYIVKNKPVLLVCKAV--PATQIFFKCNGEWVRQVDHVIERSTDGSSG 98
Db         65 DFPPRIVEHPSDLIVSKGEPATLNCKAEGRPTPTI------EWYKGGERVETDKDDPRSH 118

Qy         99 EPTME------VRINVSRQQVEKVFGLEEYWCQCVAWSSSGTTKSQKAYIRIARLRKNFE 152
Db        119 RMLLPSGSLFFLRIVHGRKSRP-----DEGVYVCVARNYLGEAVSHNASLEVAILRDDFR 173

Qy        153 QEPLAKEVSLEQGIVLPCRPPEGIPPAEVEWLRNEDLVDPSLDPNVYIT-REHSLVVRQA 211
Db        174 QNPSDVMVAVGEPAVMECQPPRGHPEPTISWKKDGSPLD---DKDERITIRGGKLMITYT 230

Qy        212 RLADTANYTCVAKNIVARRRSASAAVIVYVNGGW-------STWTEWSV---CSASCGRG 261
Db        231 RKSDAGKYVCVGTNMVGERESEVAELTVLERPSFVKRPSNLAVTVDDSAEFKCEA---RG 287

Qy        262 -------WQK------------------RSRSCTNPAPLNGGAFCEGQNVHDRTVSSLLV 296
Db        288 DPVPTVRWRKDDGELPKSRYEIRDDHTLKIRKVT--AGDMGSYTCVAENMVGKAEASATL 345

Qy        297 SVDGS----WSPWSKWSACGLDCTHWRSRECSDPAP-----RNGGEEC-------QGTDL 340
Db        346 TVQEPPHFVVKPRDQVVALGRTVT-FQCEATGNPQPAIFWRREGSQNLLFSYQPPQSSSR 404

Qy        341 DTRNCTSDLCVHSASGPEDVALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSI 400
Db        405 FSVSQTGDLTITNVQ-RSDVGYY-------ICQTL-----------------NVA-GSI 437

Qy        401 LTSGFQPVSIKPSKADNPHLLTIQPDLSTTTTYQGSL---CPRQDGPSPKF--------- 448
Db        438 ITKAY--LEVTDVIADRPPPVIRQGPVNQTVAVDGTFVLSCVATGSPVPTILWRKDGVLV 495

Qy        449 --------QLTNGHL---LSPLGG-GRHTLHHSSPTSEAEEFVSRLSTQNYFRSLPRGTS 496
Db        496 STQDSRIKQLENGVLQIRYAKLGDTGRYTCIASTPSGEA--------------------- 534
```

```
Qy        497 NMTYGTFNFLGGRLMIPNTGISLLIP----PDAIPRGKIYEIYLTLHKPE--DV---RLP 547
                |:  :          : :   |: :  |    |: ||              ||| ||     :
Db        535 --TWSAY------IEVQEFGVPVQPPRPTDPNLIPSAP--------SKPEVTDVSRNTVT 578

Qy        548 LAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGSWEQDVLHL 607
              |:    | |            |    |  :| |  |    |  ||      :     | |  : :
Db        579 LSWQPNLNS--------GATPTSYIIEAFSHA---SGSSWQTVAENVKTETSAIKGL--- 624

Qy        608 GEEAPSHLYYCQLEASACYVFTE--QLGRFALVGEALSVAAA---KRLK-------LLLF 655
                  |: :|   : |:  |  ::  |:         : |  :     |:::          | |
Db        625 ---KPNAIYLFLVRAANAYGISDPSQISDPVKTQDVLPTSQGVDHKQVQRELGNAVLHLH 681

Qy        656 APVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHD 715
                |    :|     :         | | |   : :|     | :::  |    :  :|         : :
Db        682 NPTVLSSSSIEV-----HWTVDQQSQYIQ-----GYKILYRPSGANHGES----DWLVFE 727

Qy        716 VPSSLWKSKLL------VSYQ--EIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWV 767
              | :      | ::         |:|:    ||:: :  |       :      ||| :||          :  |
Db        728 VRTPAKNSVVIPDLRKGVNYEIKARPFFNEFQGADSEIKFAKTLEE-APSAPPQG--VTV 784

Qy        768 WQVEGDGQSFSINFN-ITKDTRFAEL-------LALESEAGVPALVGPSAFK--IPFLIR 817
                : :|:| :  :::      :||:     :               |  |:    :      |   |  |     ||||:
Db        785 SKNDGNGTAILVSWQPPPEDTQNGMVQEYKVWCLGNETRYHINKTVDGSTFSVVIPFLV- 843

Qy        818 QKIISSLDPPCRRGA--DWRTLAQKLHLDSH--------------LSFFASKPSPTA-- 858
                 |   |::         ||      ::    |  :  ||:|                      :|        :|:    |
Db        844 PGIRYSVEVAASTGAGSGVKSEPQFIQLDAHGNPVSPEDQVSLAQQISDVVKQPAFIAGI 903

Qy        859 --------MILNLWEARHFP--NGNLSQLAA-------------------AVAGLGQPD 888
                         |: ::|    ||          ||     |     |                             ||:     |:|
Db        904 GAACWIILMVFSIWLYRHRKKRNGLTSTYAGIRKVPSFTFTPTVTYQRGGEAVSSGGRP- 962

Qy        889 AGLFTVSE 896
                 ||   :||
Db        963 -GLLNISE 969
```

Search completed: October 6, 2004, 18:20:40
Job time : 29 secs

```
OM protein - protein search, using sw model

Run on:         October  6, 2004, 12:45:02 ; Search time 88 Seconds
                                              (without alignments)
                                              2886.480 Million cell updates/sec

Title:          US-09-970-944-2
Perfect score:  4787
Sequence:       1 MAVRPGLWPALLGIVLAAWL..........AVAGLGQPDAGLFTVSEAEC 899

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters:      1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       A_Geneseq_29Jan04:*
                 1:  geneseqp1980s:*
                 2:  geneseqp1990s:*
                 3:  geneseqp2000s:*
                 4:  geneseqp2001s:*
                 5:  geneseqp2002s:*
                 6:  geneseqp2003as:*
                 7:  geneseqp2003bs:*
                 8:  geneseqp2004s:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 4787 | 100.0 | 899 | 5 | AAU79939 | Aau79939 Human UNC |
| 2 | 4698.5 | 98.2 | 898 | 5 | AAU85403 | Aau85403 Human pro |
| 3 | 4688.5 | 97.9 | 898 | 5 | AAU97899 | Aau97899 Human net |
| 4 | 4545.5 | 95.0 | 898 | 2 | AAW78898 | Aaw78898 Rat UNC-5 |
| 5 | 4545.5 | 95.0 | 898 | 5 | AAU10543 | Aau10543 Rat netri |
| 6 | 4545.5 | 95.0 | 898 | 5 | AAU97900 | Aau97900 Rat netri |
| 7 | 4434 | 92.6 | 943 | 4 | AAM79128 | Aam79128 Human pro |
| 8 | 4379.5 | 91.5 | 842 | 5 | AAU74818 | Aau74818 Human REP |
| 9 | 2789 | 58.3 | 556 | 2 | AAW78899 | Aaw78899 Human UNC |
| 10 | 2673.5 | 55.8 | 931 | 4 | AAB50691 | Aab50691 Human UNC |
| 11 | 2673.5 | 55.8 | 931 | 7 | ADE63098 | Ade63098 Human Pro |
| 12 | 2673.5 | 55.8 | 982 | 4 | ABG11551 | Abg11551 Novel hum |
| 13 | 2497 | 52.2 | 945 | 7 | ADE63096 | Ade63096 Rat Prote |
| 14 | 2490.5 | 52.0 | 946 | 5 | ABG61795 | Abg61795 Novel UNC |
| 15 | 2490 | 52.0 | 943 | 2 | AAW78900 | Aaw78900 Rat UNC-5 |
| 16 | 2484 | 51.9 | 933 | 5 | AAO18734 | Aao18734 Human NOV |
| 17 | 2484 | 51.9 | 933 | 5 | AAO18735 | Aao18735 Human NOV |
| 18 | 2479 | 51.8 | 945 | 4 | AAU12244 | Aau12244 Human PRO |
| 19 | 2479 | 51.8 | 945 | 6 | ABO17688 | Abo17688 Novel hum |
| 20 | 2479 | 51.8 | 945 | 6 | ABU80942 | Abu80942 Human PRO |
| 21 | 2479 | 51.8 | 945 | 6 | ABU66642 | Abu66642 Human PRO |
| 22 | 2479 | 51.8 | 945 | 6 | ABU59723 | Abu59723 Novel sec |
| 23 | 2479 | 51.8 | 945 | 6 | ABO24913 | Abo24913 Human sec |
| 24 | 2479 | 51.8 | 945 | 6 | ABU66918 | Abu66918 Human sec |
| 25 | 2479 | 51.8 | 945 | 6 | ADA45665 | Ada45665 Novel hum |
| 26 | 2479 | 51.8 | 945 | 6 | ADA76096 | Ada76096 Human PRO |
| 27 | 2479 | 51.8 | 945 | 6 | ADA18746 | Ada18746 Human PRO |
| 28 | 2479 | 51.8 | 945 | 6 | ADA61369 | Ada61369 Homo sapi |
| 29 | 2479 | 51.8 | 945 | 6 | ADB19154 | Adb19154 Novel hum |
| 30 | 2479 | 51.8 | 945 | 6 | ADB27695 | Adb27695 Human PRO |
| 31 | 2479 | 51.8 | 945 | 6 | ADA86174 | Ada86174 Novel hum |
| 32 | 2479 | 51.8 | 945 | 6 | ADB15738 | Adb15738 Human PRO |
| 33 | 2479 | 51.8 | 945 | 6 | ADA47524 | Ada47524 Human PRO |
| 34 | 2479 | 51.8 | 945 | 6 | ADA67319 | Ada67319 Human PRO |
| 35 | 2479 | 51.8 | 945 | 6 | ADB30326 | Adb30326 Human PRO |
| 36 | 2479 | 51.8 | 945 | 6 | ADA85622 | Ada85622 Novel hum |
| 37 | 2479 | 51.8 | 945 | 6 | ADA96834 | Ada96834 Human PRO |
| 38 | 2479 | 51.8 | 945 | 6 | ADA79138 | Ada79138 Human PRO |
| 39 | 2479 | 51.8 | 945 | 6 | ADA87277 | Ada87277 Novel hum |
| 40 | 2479 | 51.8 | 945 | 6 | ADB16479 | Adb16479 Human PRO |
| 41 | 2479 | 51.8 | 945 | 6 | ADA91571 | Ada91571 Novel hum |
| 42 | 2479 | 51.8 | 945 | 6 | ADB14634 | Adb14634 Human PRO |
| 43 | 2479 | 51.8 | 945 | 6 | ADB18595 | Adb18595 Novel hum |
| 44 | 2479 | 51.8 | 945 | 6 | ADA93810 | Ada93810 Human PRO |
| 45 | 2479 | 51.8 | 945 | 6 | ADB19706 | Adb19706 Novel hum |

ALIGNMENTS

```
RESULT 1
AAU79939
ID   AAU79939 standard; protein; 899 AA.
XX
AC   AAU79939;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Human UNC5-like protein NOV1.
XX
KW   Human; NOVX polypeptide; cardiomyopathy; atherosclerosis; cancer;
KW   cell signal processing; metabolic pathway modulation; cancerous tissue;
KW   antibody; diabetes; transgenic animal; UNC5-like protein; NOV1;
KW   chromosome 13.
XX
OS   Homo sapiens.
XX
PN   WO200229038-A2.
XX
PD   11-APR-2002.
XX
PF   04-OCT-2001; 2001WO-US031377.
XX
PR   04-OCT-2000; 2000US-0237862P.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Herrmann JL,  Rastelli L,  Shimkets RA;
XX
DR   WPI; 2002-340104/37.
DR   N-PSDB; ABK49422.
XX
PT   Novel isolated NOVX polypeptide, and encoded polynucleotide, useful for
PT   treating cardiomyopathy, artherosclerosis, and cancer.
XX
PS   Claim 1; Page 9; 180pp; English.
XX
CC   The present invention relates to a new NOVX polypeptide having a 900
CC   (NOV1), 4349 (NOV2), 940 (NOV3), 798 (NOV4), 865 (NOV5), or 331 (NOV6)
CC   residue amino acid sequence, as given in the specification. The novel
CC   polypeptide, and its encoding polynucleotide, are used to treat
CC   cardiomyopathy, atherosclerosis, cancer or a disease related to cell
CC   signal processing and metabolic pathway modulation, in a human. Detecting
CC   the polypeptide or polynucleotide is useful for identifying cancerous
CC   tissue. The antibody can be used to treat diabetes or cancer. The host
CC   cells can be used to produce non-human transgenic animals useful in drug
CC   screening. The present amino acid sequence is that of the human UNC5-like
CC   protein NOV1 of the invention. This sequence is encoded by the human UNC5
```

```
CC   -like NOV1 gene located on chromosome 13
XX
SQ   Sequence 899 AA;

  Query Match             100.0%;  Score 4787;  DB 5;  Length 899;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  899;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60

Qy         61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGEPTMEVRINVSRQQVEKVFGLE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGEPTMEVRINVSRQQVEKVFGLE 120

Qy        121 EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180

Qy        181 VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY 240

Qy        241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDG 300

Qy        301 SWSPWSKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCVHSASGPEDV 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 SWSPWSKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCVHSASGPEDV 360

Qy        361 ALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHL 420

Qy        421 LTIQPDLSTTTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEEFVS 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 LTIQPDLSTTTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEEFVS 480

Qy        481 RLSTQNYFRSLPRGTSNMTYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLHK 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 RLSTQNYFRSLPRGTSNMTYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLHK 540

Qy        541 PEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGSW 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 PEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGSW 600

Qy        601 EQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAAKRLKLLLFAPVAC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 EQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAAKRLKLLLFAPVAC 660

Qy        661 TSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSL 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSL 720

Qy        721 WKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSIN 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 WKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSIN 780

Qy        781 FNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQK 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 FNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQK 840

Qy        841 LHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 899
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 LHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 899

RESULT 2
AAU85403
ID   AAU85403 standard; protein; 898 AA.
XX
AC   AAU85403;
XX
DT   21-MAY-2002  (first entry)
XX
DE   Human protein NOV1.
XX
KW   Human; NOVX; cardiomyopathy; atherosclerosis; diabetes;
KW   cell signal processing disorder; metabolic disorder; obesity; infection;
KW   anorexia; cancer-associated cachexia; cancer; neurodegenerative disorder;
KW   Alzheimer's disease; Parkinson's disease; immune disorder;
KW   haematopoietic disorders; dyslipidaemia; pain; asthma; hypertension;
KW   osteoporosis; Crohn's disease; multiple sclerosis; angina pectoris;
KW   myocardial infarction; ulcer; allergy; benign prostatic hypertrophy;
KW   psychosis; neurological disorder; anxiety; schizophrenia;
KW   manic depression; dementia; dyskinesia; Huntington's disease;
KW   Gilles de la Tourette's syndrome; gene therapy.
XX
OS   Homo sapiens.
XX
PN   WO200210216-A2.
XX
PD   07-FEB-2002.
XX
PF   30-JUL-2001; 2001WO-US024225.
XX
PR   28-JUL-2000; 2000US-0221409P.
PR   04-AUG-2000; 2000US-0222840P.
PR   04-AUG-2000; 2000US-0223752P.
PR   04-AUG-2000; 2000US-0223762P.
PR   04-AUG-2000; 2000US-0223769P.
PR   04-AUG-2000; 2000US-0223770P.
PR   14-AUG-2000; 2000US-0225146P.
PR   15-AUG-2000; 2000US-0225392P.
PR   15-AUG-2000; 2000US-0225470P.
PR   16-AUG-2000; 2000US-0225697P.
PR   01-FEB-2001; 2001US-0263662P.
PR   05-APR-2001; 2001US-0281645P.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Padigaru M,  Mezes P,  Mishra V,  Burgess C,  Casman S,  Grosse WM;
PI   Alsobrook JP,  Lepley DM,  Gerlach VL,  Macdougall JR,  Smithson G;
XX
DR   WPI; 2002-180074/23.
DR   N-PSDB; ABK37922.
XX
PT   New isolated cytoplasmic, nuclear, membrane bound, or secreted
PT   polypeptide, useful for treating cardiomyopathy, atherosclerosis,
PT   infections, cancer, neurodegenerative, metabolic, hematopoietic and
PT   immune disorders.
XX
PS   Claim 1; Page 11; 213pp; English.
XX
CC   The invention relates to an isolated cytoplasmic, nuclear, membrane
CC   bound, or secreted polypeptide (NOVX, x= 1-14) their variants or mature
CC   form. Also included are the nucleic acids encoding the NOVX proteins, a
CC   vector comprising the nucleic acid, a cell comprising the vector, an anti
CC   -NOVX antibody and modulators of NOVX. NOVX, the nucleic acid and the
CC   antibody are useful for treating or preventing a NOVX-associated
CC   disorder, where the disorder is selected from cardiomyopathy,
CC   atherosclerosis, diabetes, a disorder related to cell signal processing
CC   and metabolic pathway modulation, metabolic disorders, obesity,
CC   infectious disease, anorexia, cancer-associated cachexia, cancer,
CC   neurodegenerative disorders, Alzheimer's disease, Parkinson's disease,
CC   immune disorders, haematopoietic disorders, and the various
CC   dyslipidaemias, metabolic disturbances associated with obesity, the
CC   metabolic syndrome X and wasting disorders associated with chronic
CC   diseases, bacterial, fungal, protozoal and viral infections, pain,
CC   bulimia, asthma, hypertension, urinary retention, osteoporosis, Crohn's
CC   disease, multiple sclerosis, Albright Hereditary Osteodystrophy, angina
CC   pectoris, myocardial infarction, ulcer, allergy, benign prostatic
CC   hypertrophy, and psychotic and neurological disorders, including anxiety,
CC   schizophrenia, manic depression, delirium, dementia, and dyskinesias,
```

```
CC   such as Huntington's disease and Gilles de la Tourette's syndrome. The
CC   nucleic acid is useful in gene therapy. The present sequence represents a
CC   NOVX protein
XX
SQ   Sequence 898 AA;

  Query Match             98.2%;  Score 4698.5;  DB 5;  Length 898;
  Best Local Similarity   98.7%;  Pred. No. 0;
  Matches  888;  Conservative    2;  Mismatches    7;  Indels    3;  Gaps    3;

Qy          1 MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60

Qy         61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGEPTMEVRINVSRQQVEKVFGLE 120
              ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db         61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSRQQVEKVFGLE 120

Qy        121 EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180

Qy        181 VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY 240

Qy        241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDG 300
              ||||||||||||||||||||||||||||||||||||||||||||||   :|  :  | |||
Db        241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNV-QKTACATLCPVDG 299

Qy        301 SWSPWSKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCVHSASGPEDV 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        300 SWSPWSKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCVHSASGPEDV 359

Qy        361 ALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        360 ALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHL 419

Qy        421 LTIQPDLS-TTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEEFV 479
              |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        420 LTIQPDLSTTTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEEFV 479

Qy        480 SRLSTQNYFRSLPRGTSNMTYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLH 539
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 SRLSTQNYFRSLPRGTSNMTYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLH 539

Qy        540 KPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGS 599
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 KPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGS 599

Qy        600 WEQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAAKRLKLLLFAPVA 659
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 WE-DVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAAKRLKLLLFAPVA 658

Qy        660 CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSS 719
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        659 CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSS 718

Qy        720 LWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSI 779
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        719 LWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSI 778

Qy        780 NFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQ 839
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        779 NFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQ 838

Qy        840 KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 899
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        839 KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 898


RESULT 3
AAU97899
ID   AAU97899 standard; protein; 898 AA.
XX
AC   AAU97899;
XX
DT   27-AUG-2002  (first entry)
XX
DE   Human netrin binding membrane receptor UNC5H-1 protein.
XX
KW   Netrin binding membrane receptor; receptor; UNC5H-1; human; nootropic;
KW   neuroprotective; cytostatic; antiparkinsonian; cerebroprotective; cancer;
KW   central nervous system; CNS; stroke; Parkinson's disease;
KW   multiple sclerosis; Alzheimer's disease.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Domain          152. .223
FT                   /note= "Immunoglobulin domain "
FT   Domain          247. .294
FT                   /note= "Thrombospondine type 1 domain "
FT   Domain          302. .348
FT                   /note= "Thrombospondine type 1 domain"
FT   Region          361. .382
FT                   /note= "Transmembrane region"
FT   Domain          495. .598
FT                   /note= "ZU5 domain"
FT   Domain          817. .897
FT                   /note= "Death domain"
XX
PN   WO200233080-A2.
XX
PD   25-APR-2002.
XX
PF   15-OCT-2001; 2001WO-EP011891.
XX
PR   16-OCT-2000; 2000US-0240061P.
XX
PA   (FARB ) BAYER AG.
XX
PI   Koehler RH;
XX
DR   WPI; 2002-463314/49.
DR   N-PSDB; ABK52891.
XX
PT   Novel human netrin binding membrane receptor polypeptide and
PT   polynucleotides for identifying modulating agents useful in treating
PT   diseases e.g. Parkinson's disease, multiple sclerosis, stroke,
PT   Alzheimer's disease.
XX
PS   Claim 1; Fig 2; 94pp; English.
XX
CC   This invention relates to the DNA and protein sequences of a novel
CC   purified human netrin binding membrane receptor, UNC5H-1. The DNA
CC   sequence of the invention is useful as a probe for detecting a nucleic
CC   acid encoding the UNC5H-1 protein in a biological sample. The sequences
CC   of the invention are useful to screen for agents which decrease the
CC   activity of the UNC5H-1 protein. The sequences are also useful for
CC   screening agents which regulate (modulate) the activity of the protein of
CC   the invention. A pharmaceutical composition containing the protein of the
CC   invention or a reagent that modulates the activity of the UNC5H-1 protein
CC   may be useful for treating a UNC5H-1 dysfunction related disease such as
CC   cancer or a central nervous system (CNS) disorders (e.g, Parkinson's
CC   disease, multiple sclerosis, stroke and Alzheimer's disease). Fusion
CC   proteins comprising the UNC5H-1 protein are useful for generating
CC   antibodies and for in various assay systems, and the protein can be used
CC   as a bait protein in a two-hybrid assay or three-hybrid assay. The method
CC   of the invention is useful for detecting a coding sequence for the UNC5H-
CC   1 protein. The present sequence represents the human netrin binding
CC   membrane receptor UNC5H-1 protein of the invention
XX
SQ   Sequence 898 AA;

  Query Match             97.9%;  Score 4688.5;  DB 5;  Length 898;
```

```
  Best Local Similarity   98.4%;  Pred. No. 0;
  Matches  886;  Conservative    3;  Mismatches    8;  Indels    3;  Gaps    3;

Qy          1 MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60

Qy         61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGEPTMEVRINVSRQQVEKVFGLE 120
              |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db         61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSRQQVEKVFGLE 120

Qy        121 EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180
              |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db        121 EYWCQCVAWSSSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180

Qy        181 VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRSASAAVIVY 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRSASAAVIVY 240

Qy        241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDG 300
              ||||||||||||||||||||||||||||||||||||||||||||||  :|  :|  |||
Db        241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNV-QKTACATLCPVDG 299

Qy        301 SWSPWSKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCVHSASGPEDV 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db        300 SWSPWSKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCVHTASGPEDV 359

Qy        361 ALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        360 ALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHL 419

Qy        421 LTIQPDLS-TTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEEFV 479
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        420 LTIQPDLSTTTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEEFV 479

Qy        480 SRLSTQNYFRSLPRGTSNMTYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLH 539
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 SRLSTQNYFRSLPRGTSNMTYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLH 539

Qy        540 KPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGS 599
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 KPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGS 599

Qy        600 WEQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAAKRLKLLLFAPVA 659
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 WE-DVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAAKRLKLLLFAPVA 658

Qy        660 CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSS 719
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        659 CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSS 718

Qy        720 LWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSI 779
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        719 LWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSI 778

Qy        780 NFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQ 839
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        779 NFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQ 838

Qy        840 KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 899
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        839 KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 898
```

```
RESULT 4
AAW78898
ID   AAW78898 standard; protein; 898 AA.
XX
AC   AAW78898;
XX
DT   25-MAR-2003  (revised)
DT   21-DEC-1998  (first entry)
XX
DE   Rat UNC-5 homologue UNC5H-1.
XX
KW   UNC-5; UNC5H-1; rat; netrin receptor; cell migration; axon guidance;
KW   diagnosis; therapy.
XX
OS   Rattus sp.
XX
FH   Key             Location/Qualifiers
FT   Peptide         580..594
FT                   /note= "peptide used to raise rabbit polyclonal antisera"
XX
PN   WO9837085-A1.
XX
PD   27-AUG-1998.
XX
PF   19-FEB-1998;   98WO-US003143.
XX
PR   19-FEB-1997;   97US-00808982.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Tessier-Lavigne M,  Leonardo ED,  Hinck L,  Masu M,  Keinomasu K;
XX
DR   WPI; 1998-495364/42.
DR   N-PSDB; AAV52940.
XX
PT   Netrin-binding, vertebrate proteins - useful for diagnosis, therapy and
PT   the biopharmaceutical industry.
XX
PS   Claim 1; Page 19-22; 32pp; English.
XX
CC   UNC5H-1 and UNC5H-2 (see AAW78900) are rat homologues of Caenorhabditis
CC   elegans UNC-5 protein. Their amino acid sequences were deduced from
CC   isolated unc5h cDNA clones (see AAV52940 and AAV52942) isolated from an
CC   E18 brain cDNA library. The predicted proteins show similarity with UNC-
CC   5, possess 2 predicted Ig-like domains and 2 predicted thrombospondin
CC   type-1 repeats, a predicted membrane spanning region, and a large
CC   intracellular domain. They are predicted to be involved in cell migration
CC   and axon guidance, and are characterised as receptor proteins for
CC   netrins. Human UNC5H-1 (see AAW78899) and UNC5H-2 (see AAW78901) proteins
CC   are also claimed. Vertebrate UNC-5 proteins may be produced recombinantly
CC   from transfected host cells. The invention also provides unc-5
CC   hybridisation probes and primers, vertebrate UNC-5-specific binding
CC   agents such as specific antibodies, and methods of making and using the
CC   subject compositions in diagnosis (e.g. genetic hybridisation screens for
CC   vertebrate unc-5 transcripts), therapy (e.g. gene therapy to modulate
CC   vertebrate unc-5 gene expression) and in the biopharmaceutical industry
CC   (e.g. as immunogens, reagents for modulating cell guidance, reagents for
CC   screening chemical libraries for lead pharmacological agents, etc.).
CC   (Updated on 25-MAR-2003 to correct PI field.)
XX
SQ   Sequence 898 AA;

  Query Match             95.0%;  Score 4545.5;  DB 2;  Length 898;
  Best Local Similarity   94.7%;  Pred. No. 0;
  Matches  852;  Conservative   19;  Mismatches   26;  Indels    3;  Gaps    3;

Qy          1 MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60
              ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAVRPGLWPVLLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60

Qy         61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGEPTMEVRINVSRQQVEKVFGLE 120
              |||||||||||||||||||||||||||||||||| ||| |||||||||||||||||||||
Db         61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDSSSGLPTMEVRINVSRQQVEKVFGLE 120

Qy        121 EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180
              |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db        121 EYWCQCVAWSSSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180

Qy        181 VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRSASAAVIVY 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        181 VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRSTSAAVIVY 240
```

```
Qy        241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDG 300
              ||||||||||||||||||||||||||||||||||||||||||||||  :|    |  |||
Db        241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNV-QKTACATLCPVDG 299

Qy        301 SWSPWSKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCVHSASGPEDV 360
              ||| |||||||||||||||||||||||||||||||:| ||||||||||||:|:|| ||||
Db        300 SWSSWSKWSACGLDCTHWRSRECSDPAPRNGGEECRGADLDTRNCTSDLCLHTASCPEDV 359

Qy        361 ALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHL 420
              |||:||:|||||| |||   :: |||||||||||||||||||||||||||||||||||||
Db        360 ALYIGLVAVAVCLFLLLALGLIYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHL 419

Qy        421 LTIQPDLS-TTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEEFV 479
              |||||||| |||||||||| |||||||||||:||||||||| |||||||||||||||:||
Db        420 LTIQPDLSTTTTTYQGSLCSRQDGPSPKFQLSNGHLLSPLGSGRHTLHHSSPTSEAEDFV 479

Qy        480 SRLSTQNYFRSLPRGTSNMTYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLH 539
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        480 SRLSTQNYFRSLPRGTSNMAYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLH 539

Qy        540 KPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGS 599
              ||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||
Db        540 KPEDVRLPLAGCQTLLSPVVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGS 599

Qy        600 WEQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAAKRLKLLLFAPVA 659
              || ||||||||:||||||||||| ||||||||||||||||||||||| |||:||||||||
Db        600 WE-DVLHLGEESPSHLYYCQLEAGACYVFTEQLGRFALVGEALSVAATKRLRLLLFAPVA 658

Qy        660 CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSS 719
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        659 CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSS 718

Qy        720 LWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSI 779
              |||||||||||||||||||||||:||||||||||:: ||||||||:||||||||||||:|
Db        719 LWKSKLLVSYQEIPFYHIWNGTQQYLHCTFTLERINASTSDLACKVWVWQVEGDGQSFNI 778

Qy        780 NFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQ 839
              ||||||||||||||||||| ||||||||||||||||||||||:|||||| ||||||||||
Db        779 NFNITKDTRFAELLALESEGGVPALVGPSAFKIPFLIRQKIIASLDPPCSRGADWRTLAQ 838

Qy        840 KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 899
              ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        839 KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLGQLAAAVAGLGQPDAGLFTVSEAEC 898
```

RESULT 5
AAU10543
ID   AAU10543 standard; protein; 898 AA.
XX
AC   AAU10543;
XX
DT   14-FEB-2002  (first entry)
XX
DE   Rat netrin receptor UNC5H1 (YSG7) polypeptide.
XX
KW   YSG; YSG7; schizophrenia; chronic animal model; LCGU; netrin receptor;
KW   local cerebral glucose utilisation; phosphodiesterase 1-alpha; UNC5H1;
KW   calcium-independent alpha-latrotoxin receptor; CIRL; trkE; synapsin 1A;
KW   epithelial discoidin domain receptor 1; synapsin 1B; neuroleptic;
KW   tumour necrosis factor alpha; TNF-alpha; rat.
XX
OS   Rattus sp.
XX
PN   WO200175440-A2.
XX
PD   11-OCT-2001.
XX
PF   02-APR-2001; 2001WO-GB001486.
XX
PR   31-MAR-2000; 2000GB-00007880.
PR   26-MAY-2000; 2000GB-00012768.
XX
PA   (WELF-) WELFIDE CORP.
XX
PI   Cochran S,  Paterson G,  Ohashi Y,  Morris B,  Pratt J;
XX
DR   WPI; 2002-010813/01.
DR   N-PSDB; AAS16843.
XX
PT   Novel chronic animal model of schizophrenia, useful for identifying anti-
PT   psychotic drugs and genes that are associated with schizophrenia.
XX
PS   Disclosure; Fig 8b; 79pp; English.
XX
CC   The invention relates to YSG polynucleotide fragments for use in
CC   diagnosing and/or developing treatments for schizophrenia using chronic
CC   animal models. The polynucleotides and their encoded polypeptides are
CC   used for identification of compounds which modulate the expression of YSG
CC   molecules, leading to the manufacture of schizophrenia medicaments. The
CC   sequences can also be used for testing candidate compounds for any effect
CC   on the polypeptides. Anti-schizophrenic effects of a compound can be
CC   determined by measuring local cerebral glucose utilisation (LCGU) or
CC   comparing its expression level with that of a control group. The
CC   sequences are useful in the identification of genes associated with
CC   schizophrenic states and in the development of an antibody. The sequences
CC   of the invention include phosphodiesterase 1-alpha, calcium-independent
CC   alpha-latrotoxin receptors (CIRL)-1,2&3, epithelial discoidin domain
CC   receptor 1 (trkE), netrin receptor (UNC5H1), synapsins 1A and AB and
CC   tumour necrosis factor (TNF) alpha. This sequence represents rat netrin
CC   receptor UNC5H1 (YSG7) polypeptide
XX
SQ   Sequence 898 AA;

```
  Query Match             95.0%;  Score 4545.5;  DB 5;  Length 898;
  Best Local Similarity   94.7%;  Pred. No. 0;
  Matches  852;  Conservative   19;  Mismatches   26;  Indels    3;  Gaps    3;

Qy          1 MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60
              ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAVRPGLWPVLLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60

Qy         61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGEPTMEVRINVSRQQVEKVFGLE 120
              |||||||||||||||||||||||||||||||||| ||| |||||||||||||||||||||
Db         61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDSSSGLPTMEVRINVSRQQVEKVFGLE 120

Qy        121 EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180
              ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        121 EYWCQCVAWSSSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180

Qy        181 VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        181 VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSTSAAVIVY 240

Qy        241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDG 300
              ||||||||||||||||||||||||||||||||||||||||||||||  :|    |  |||
Db        241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNV-QKTACATLCPVDG 299

Qy        301 SWSPWSKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCVHSASGPEDV 360
              ||| |||||||||||||||||||||||||||||||:| ||||||||||||:|:|| ||||
Db        300 SWSSWSKWSACGLDCTHWRSRECSDPAPRNGGEECRGADLDTRNCTSDLCLHTASCPEDV 359

Qy        361 ALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHL 420
              |||:||:|||||| |||   :: |||||||||||||||||||||||||||||||||||||
Db        360 ALYIGLVAVAVCLFLLLALGLIYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHL 419

Qy        421 LTIQPDLS-TTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEEFV 479
              |||||||| |||||||||| |||||||||||:||||||||| |||||||||||||||:||
Db        420 LTIQPDLSTTTTTYQGSLCSRQDGPSPKFQLSNGHLLSPLGSGRHTLHHSSPTSEAEDFV 479

Qy        480 SRLSTQNYFRSLPRGTSNMTYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLH 539
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        480 SRLSTQNYFRSLPRGTSNMAYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLH 539

Qy        540 KPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGS 599
              ||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||
Db        540 KPEDVRLPLAGCQTLLSPVVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGS 599
```

```
Qy       600 WEQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAAKRLKLLLFAPVA 659
Db       600 WE-DVLHLGEESPSHLYYCQLEAGACYVFTEQLGRFALVGEALSVAATKRLRLLLFAPVA 658

Qy       660 CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSS 719
Db       659 CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSS 718

Qy       720 LWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSI 779
Db       719 LWKSKLLVSYQEIPFYHIWNGTQQYLHCTFTLERINASTSDLACKVWVWQVEGDGQSFNI 778

Qy       780 NFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQ 839
Db       779 NFNITKDTRFAELLALESEGGVPALVGPSAFKIPFLIRQKIIASLDPPCSRGADWRTLAQ 838

Qy       840 KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 899
Db       839 KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLGQLAAAVAGLGQPDAGLFTVSEAEC 898
```

```
RESULT 6
AAU97900
ID   AAU97900 standard; protein; 898 AA.
XX
AC   AAU97900;
XX
DT   27-AUG-2002  (first entry)
XX
DE   Rat netrin binding membrane receptor UNC5H-1 protein.
XX
KW   Netrin binding membrane receptor; receptor; UNC5H-1; Rat; nootropic;
KW   neuroprotective; cytostatic; antiparkinsonian; cerebroprotective; cancer;
KW   central nervous system; CNS; stroke; Parkinson's disease;
KW   multiple sclerosis; Alzheimer's disease.
XX
OS   Rattus sp.
XX
FH   Key             Location/Qualifiers
FT   Domain          152. .223
FT                   /note= "Immunoglobulin domain "
FT   Domain          247. .294
FT                   /note= "Thrombospondine type 1 domain "
FT   Domain          302. .348
FT                   /note= "Thrombospondine type 1 domain"
FT   Region          361. .382
FT                   /note= "Transmembrane region"
FT   Domain          495. .598
FT                   /note= "ZU5 domain"
FT   Domain          817. .897
FT                   /note= "Death domain"
XX
PN   WO200233080-A2.
XX
PD   25-APR-2002.
XX
PF   15-OCT-2001; 2001WO-EP011891.
XX
PR   16-OCT-2000; 2000US-0240061P.
XX
PA   (FARB ) BAYER AG.
XX
PI   Koehler RH;
XX
DR   WPI; 2002-463314/49.
XX
PT   Novel human netrin binding membrane receptor polypeptide and
PT   polynucleotides for identifying modulating agents useful in treating
PT   diseases e.g. Parkinson's disease, multiple sclerosis, stroke,
PT   Alzheimer's disease.
XX
PS   Disclosure; Fig 3; 94pp; English.
XX
CC   This invention relates to the DNA and protein sequences of a novel
CC   purified human netrin binding membrane receptor, UNC5H-1. The DNA
CC   sequence of the invention is useful as a probe for detecting a nucleic
CC   acid encoding the UNC5H-1 protein in a biological sample. The sequences
CC   of the invention are useful to screen for agents which decrease the
CC   activity of the UNC5H-1 protein. The sequences are also useful for
CC   screening agents which regulate (modulate) the activity of the protein of
CC   the invention. A pharmaceutical composition containing the protein of the
CC   invention or a reagent that modulates the activity of the UNC5H-1 protein
CC   may be useful for treating a UNC5H-1 dysfunction related disease such as
CC   cancer or a central nervous system (CNS) disorders (e.g, Parkinson's
CC   disease, multiple sclerosis, stroke and Alzheimer's disease). Fusion
CC   proteins comprising the UNC5H-1 protein are useful for generating
CC   antibodies and for in various assay systems, and the protein can be used
CC   as a bait protein in a two-hybrid assay or three-hybrid assay. The method
CC   of the invention is useful for detecting a coding sequence for the UNC5H-
CC   1 protein. The present sequence represents the Rat netrin binding
CC   membrane receptor UNC5H-1 protein of the invention
XX
SQ   Sequence 898 AA;

  Query Match             95.0%;  Score 4545.5;  DB 5;  Length 898;
  Best Local Similarity   94.7%;  Pred. No. 0;
  Matches  852;  Conservative   19;  Mismatches   26;  Indels    3;  Gaps    3;

Qy         1 MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60
Db         1 MAVRPGLWPVLLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60

Qy        61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGEPTMEVRINVSRQQVEKVFGLE 120
Db        61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDSSSGLPTMEVRINVSRQQVEKVFGLE 120

Qy       121 EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180
Db       121 EYWCQCVAWSSSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180

Qy       181 VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY 240
Db       181 VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSTSAAVIVY 240

Qy       241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDG 300
Db       241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNV-QKTACATLCPVDG 299

Qy       301 SWSPWSKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCVHSASGPEDV 360
Db       300 SWSSWSKWSACGLDCTHWRSRECSDPAPRNGGEECRGADLDTRNCTSDLCLHTASCPEDV 359

Qy       361 ALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHL 420
Db       360 ALYIGLVAVAVCLFLLLLALGLIYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHL 419

Qy       421 LTIQPDLS-TTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEEFV 479
Db       420 LTIQPDLSTTTTTYQGSLCSRQDGPSPKFQLSNGHLLSPLGSGRHTLHHSSPTSEAEDFV 479

Qy       480 SRLSTQNYFRSLPRGTSNMTYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLH 539
Db       480 SRLSTQNYFRSLPRGTSNMAYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLH 539

Qy       540 KPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGS 599
Db       540 KPEDVRLPLAGCQTLLSPVVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGS 599

Qy       600 WEQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAAKRLKLLLFAPVA 659
Db       600 WE-DVLHLGEESPSHLYYCQLEAGACYVFTEQLGRFALVGEALSVAATKRLRLLLFAPVA 658

Qy       660 CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSS 719
Db       659 CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSS 718
```

```
Qy        720 LWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSI 779
              |||||||||||||||||||||||:||||||||||:: ||||||||:||||||||||||:|
Db        719 LWKSKLLVSYQEIPFYHIWNGTQQYLHCTFTLERINASTSDLACKVWVWQVEGDGQSFNI 778

Qy        780 NFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQ 839
              ||||||||||||||||||| ||||||||||||||||||||||:|||||| ||||||||||
Db        779 NFNITKDTRFAELLALESEGGVPALVGPSAFKIPFLIRQKIIASLDPPCSRGADWRTLAQ 838

Qy        840 KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 899
              ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        839 KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLGQLAAAVAGLGQPDAGLFTVSEAEC 898
```

```
RESULT 7
AAM79128
ID   AAM79128 standard; protein; 943 AA.
XX
AC   AAM79128;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human protein SEQ ID NO 1790.
XX
KW   Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW   vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW   tissue growth factor; immunomodulatory; cancer; leukaemia;
KW   nervous system disorder; arthritis; inflammation.
XX
OS   Homo sapiens.
XX
PN   WO200157190-A2.
XX
PD   09-AUG-2001.
XX
PF   05-FEB-2001; 2001WO-US004098.
XX
PR   03-FEB-2000; 2000US-00496914.
PR   27-APR-2000; 2000US-00560875.
PR   20-JUN-2000; 2000US-00598075.
PR   19-JUL-2000; 2000US-00620325.
PR   01-SEP-2000; 2000US-00654936.
PR   15-SEP-2000; 2000US-00663561.
PR   20-OCT-2000; 2000US-00693325.
PR   30-NOV-2000; 2000US-00728422.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Drmanac RT,  Asundi V,  Zhou P,  Xu C,  Cao Y;
PI   Ma Y,  Zhao QA,  Wang D,  Wang J,  Zhang J,  Ren F,  Chen R,  Wang ZW;
PI   Xue AJ,  Yang Y,  Wejhrman T,  Goodrich R;
XX
DR   WPI; 2001-476283/51.
DR   N-PSDB; AAK52261.
XX
PT   Nucleic acids encoding polypeptides with cytokine-like activities, useful
PT   in diagnosis and gene therapy.
XX
PS   Claim 20; Page 4148-4150; 6221pp; English.
XX
CC   The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC   encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
CC   cytokine, cell proliferation or cell differentiation or which may induce
CC   production of other cytokines in other cell populations. The
CC   polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC   peptide therapy. The polypeptides have various cytokine-like activities,
CC   e.g. stem cell growth factor activity, haematopoiesis regulating
CC   activity, tissue growth factor activity, immunomodulatory activity and
CC   activin/inhibin activity and may be useful in the diagnosis and/or
CC   treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC   inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
CC   (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
CC   sequence listing were missing at the time of publication
XX
SQ   Sequence 943 AA;

  Query Match             92.6%;  Score 4434;  DB 4;  Length 943;
  Best Local Similarity   90.4%;  Pred. No. 0;
  Matches  853;  Conservative    4;  Mismatches   17;  Indels   70;  Gaps    7;
```

```
Qy          1 MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60
              |  || |        :|    |  ||||||||||||||||||||||||||||||||||||
Db         25 MTRRPSL--------MAGRQHGWSAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 76

Qy         61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGEPTMEVRINVSRQQVEKVFGLE 120
              |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db         77 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSRQQVEKVFGLE 136

Qy        121 EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180
              ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        137 EYWCQCVAWSSSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 196

Qy        181 VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        197 VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY 256

Qy        241 -----------------------------VNGGWSTWTEWSVCSASCGRGWQKRSRSCTN 271
                                           |||||||||||||||||||||||||||||||
Db        257 GGPRDSLVTGRGTAVPLGSDMWLSFSVRPVNGGWSTWTEWSVCSASCGRGWQKRSRSCTN 316

Qy        272 PAPLNGGAFCEGQNVHDRTVSSLLVSVDGSWSPWSKWSACGLDCTHWRSRECSDPAPRNG 331
              |||||||||||||||  :|  : |  |||||||||||||||||||||||||||||||||
Db        317 PAPLNGGAFCEGQNV-QKTACATLCPVDGSWSPWSKWSACGLDCTHWRSRECSDPAPRNG 375

Qy        332 GEECQGTDLDTRNCTSDLCVH--------------SASGPEDVALYVGLIAVAVCLVLL 376
              |||||||||||||||||||||              :|||||||||||||||||||||||
Db        376 GEECQGTDLDTRNCTSDLCVHNSYTPAPTKAMLSPAAASGPEDVALYVGLIAVAVCLVLL 435

Qy        377 LLVLILVYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHLLTIQPDLS-TTTTYQG 435
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        436 LLVLILVYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHLLTIQPDLSTTTTTYQG 495

Qy        436 SLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEEFVSRLSTQNYFRSLPRGT 495
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        496 SLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEEFVSRLSTQNYFRSLPRGT 555

Qy        496 SNMTYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLHKPEDVRLPLAGCQTLL 555
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        556 SNMTYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLHKPED------------ 603

Qy        556 SPIVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGSWEQDVLHLGEEAPSHL 615
                 ||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db        604 ---VSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGSWE-DVLHLGEEAPSHL 659

Qy        616 YYCQLEASACYVFTEQLGRFALVGEALSVAAAKRLKLLLFAPVACTSLEYNIRVYCLHDT 675
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 YYCQLEASACYVFTEQLGRFALVGEALSVAAAKRLKLLLFAPVACTSLEYNIRVYCLHDT 719

Qy        676 HDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWKSKLLVSYQEIPFY 735
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 HDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWKSKLLVSYQEIPFY 779

Qy        736 HIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLAL 795
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 HIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLAL 839

Qy        796 ESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLDSHLSFFASKPS 855
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        840 ESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLDSHLSFFASKPS 899

Qy        856 PTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 899
              ||||||||||||||||||||||||||||||||||||||||||||
Db        900 PTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 943
```

RESULT 8

```
AAU74818
ID   AAU74818 standard; protein; 842 AA.
XX
AC   AAU74818;
XX
DT   23-APR-2002  (first entry)
XX
DE   Human REPTR 1 protein.
XX
KW   REPTR; human; antiinflammatory; cytostatic; immunosuppressive; antiviral;
KW   anti-HIV; antiarthritic; anticonvulsant; nootropic; neuroprotective;
KW   antiallergic; antibody; immunogen; endometriosis;
KW   gastrointestinal disorder; gastritis; oesophageal carcinoma;
KW   Crohn's disease; irritable bowel syndrome; ulcerative colitis;
KW   endocrine disorder; hypothalamus disorder; Kallman's disease;
KW   autoimmune disease; inflammatory disease; infertility; receptor;
KW   acquired immune deficiency syndrome; AIDS; rheumatoid arthritis; allergy;
KW   osteoarthritis; diabetes mellitus; multiple sclerosis;
KW   systemic lupus erythematosus; cell proliferative disorder; cancer;
KW   developmental disorder; Duchenne muscular dystrophy;
KW   Becker muscular dystrophy; neurological disorder; epilepsy;
KW   Alzheimer's disease; Huntington's disease; reproductive disorder.
XX
OS   Homo sapiens.
XX
PN   WO200198354-A2.
XX
PD   27-DEC-2001.
XX
PF   21-JUN-2001; 2001WO-US019942.
XX
PR   21-JUN-2000; 2000US-0214027P.
PR   25-AUG-2000; 2000US-0228045P.
PR   12-DEC-2000; 2000US-0255104P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Griffin JA,  Kallick DA,  Tribouley CM,  Yue H,  Nguyen DB,  Tang YT;
PI   Lal P,  Policky JL,  Azimzai Y,  Lu DAM,  Graul R,  Yao MG,  Burford N;
PI   Hafalia AJA,  Baughn MR,  Bandman O,  Patterson C,  Yang J,  Xu Y;
PI   Gandhi AR,  Warren BA,  Ding L,  Sanjanwala MS,  Duggan BM,  Lu Y;
XX
DR   WPI; 2002-090432/12.
DR   N-PSDB; ABK15169.
XX
PT   Twelve human receptors (referred to as REPTR-1 to REPTR-12), useful in
PT   the diagnosis, treatment and prevention of gastrointestinal (e.g.
PT   gastritis), autoimmune/inflammatory (e.g. osteoarthritis) and cell
PT   proliferative (e.g. cancer)disorders.
XX
PS   Claim 45; Page 111-113; 157pp; English.
XX
CC   This invention relates to twelve human receptors cDNA sequences referred
CC   to as REPTR-1 to REPTR-12), and the proteins encoded thereby. The
CC   proteins of the invention may have antiinflammatory, cytostatic,
CC   immunosuppressive, antiviral, anti-HIV, antiarthritic, muscular active
CC   general, anticonvulsant, nootropic, neuroprotective, antiallergic
CC   activities. The sequences of the invention may be used to produce REPTR
CC   agonists or antagonists, and the protein sequences may be used to raise
CC   anti-REPTR antibodies. These molecules and the REPTR polynucleotides and
CC   polypeptides of the invention are useful in the diagnosis,treatment and
CC   prevention of gastrointestinal (e.g. gastritis, oesophageal carcinoma,
CC   Crohn's disease, irritable bowel syndrome, ulcerative colitis), endocrine
CC   (e.g. hypothalamus disorder, Kallman's disease), autoimmune/ inflammatory
CC   (e.g. acquired immune deficiency syndrome (AIDS), rheumatoid arthritis,
CC   allergies, osteoarthritis, diabetes mellitus, multiple sclerosis,
CC   systemic lupus erythematosus), cell proliferative (e.g. cancer),
CC   developmental (e.g. Duchenne and Becker muscular dystrophy), neurological
CC   (e.g. epilepsy, Alzheimer's disease, Huntington's disease) and
CC   reproductive (e.g. infertility, endometriosis) disorders. Numerous other
CC   examples of each disorder are given in the specification. The present
CC   sequence represents the human REPTR1 protein sequence of the invention
XX
SQ   Sequence 842 AA;

  Query Match             91.5%;  Score 4379.5;  DB 5;  Length 842;
  Best Local Similarity   93.1%;  Pred. No. 0;
  Matches  838;  Conservative    1;  Mismatches    2;  Indels   59;  Gaps    3;

Qy          1 MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60

Qy         61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGEPTMEVRINVSRQQVEKVFGLE 120
              |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db         61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSRQQVEKVFGLE 120

Qy        121 EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180
              ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        121 EYWCQCVAWSSSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180

Qy        181 VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY 240

Qy        241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDG 300
                                                                       |||
Db        241 ---------------------------------------------------------VDG 243

Qy        301 SWSPWSKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCVHSASGPEDV 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db        244 SWSPWSKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCVHTASGPEDV 303

Qy        361 ALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 ALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHL 363

Qy        421 LTIQPDLS-TTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEEFV 479
              |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 LTIQPDLSTTTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEEFV 423

Qy        480 SRLSTQNYFRSLPRGTSNMTYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLH 539
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 SRLSTQNYFRSLPRGTSNMTYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLH 483

Qy        540 KPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGS 599
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 KPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGS 543

Qy        600 WEQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAAKRLKLLLFAPVA 659
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        544 WE-DVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAAKRLKLLLFAPVA 602

Qy        660 CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSS 719
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        603 CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSS 662

Qy        720 LWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSI 779
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        663 LWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSI 722

Qy        780 NFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQ 839
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        723 NFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQ 782

Qy        840 KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 899
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        783 KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 842

RESULT 9
AAW78899
ID   AAW78899 standard; protein; 556 AA.
XX
AC   AAW78899;
XX
```

```
DT   25-MAR-2003  (revised)
DT   21-DEC-1998  (first entry)
XX
DE   Human UNC-5 homologue UNC5H-1.
XX
KW   UNC-5; UNC5H-1; human; netrin receptor; cell migration; axon guidance;
KW   diagnosis; therapy.
XX
OS   Homo sapiens.
XX
FH   Key                Location/Qualifiers
FT   Misc-difference    7
FT                      /note= "encoded by TG"
FT   Misc-difference    67
FT                      /note= "encoded by ATCT"
FT   Misc-difference    256
FT                      /note= "encoded by GC"
FT   Misc-difference    262
FT                      /note= "encoded by TG"
FT   Misc-difference    360
FT                      /note= "encoded by AG"
FT   Misc-difference    367
FT                      /note= "encoded by CC"
FT   Misc-difference    370
FT                      /note= "encoded by TC"
FT   Misc-difference    542
FT                      /note= "encoded by GG"
XX
PN   WO9837085-A1.
XX
PD   27-AUG-1998.
XX
PF   19-FEB-1998;   98WO-US003143.
XX
PR   19-FEB-1997;   97US-00808982.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Tessier-Lavigne M,  Leonardo ED,  Hinck L,  Masu M,  Keinomasu K;
XX
DR   WPI; 1998-495364/42.
DR   N-PSDB; AAW78899.
XX
PT   Netrin-binding, vertebrate proteins - useful for diagnosis, therapy and
PT   the biopharmaceutical industry.
XX
PS   Claim 1; Page 22-23; 32pp; English.
XX
CC   UNC5H-1 and UNC5H-2 (see AAW78901) are human homologues of Caenorhabditis
CC   elegans UNC-5 protein. Their amino acid sequences were deduced from
CC   isolated unc5h cDNA clones (see AAV52941 and AAV52943) isolated from an
CC   embryonic brain cDNA library. The predicted proteins show similarity with
CC   UNC-5, possess 2 predicted Ig-like domains and 2 predicted thrombospondin
CC   type-1 repeats, a predicted membrane spanning region, and a large
CC   intracellular domain. They are predicted to be involved in cell migration
CC   and axon guidance, and are characterised as receptor proteins for
CC   netrins. Rat UNC5H-1 (see AAW78898) and UNC5H-2 (see AAW78900) proteins
CC   are also claimed. Vertebrate UNC-5 proteins may be produced recombinantly
CC   from transfected host cells. The invention also provides unc-5
CC   hybridisation probes and primers, vertebrate UNC-5-specific binding
CC   agents such as specific antibodies, and methods of making and using the
CC   subject compositions in diagnosis (e.g. genetic hybridisation screens for
CC   vertebrate unc-5 transcripts), therapy (e.g. gene therapy to modulate
CC   vertebrate unc-5 gene expression) and in the biopharmaceutical industry
CC   (e.g. as immunogens, reagents for modulating cell guidance, reagents for
CC   screening chemical libraries for lead pharmacological agents, etc.).
CC   (Updated on 25-MAR-2003 to correct PI field.)
XX
SQ   Sequence 556 AA;

  Query Match             58.3%;  Score 2789;  DB 2;  Length 556;
  Best Local Similarity   96.6%;  Pred. No. 4.5e-233;
  Matches  538;  Conservative    1;  Mismatches   16;  Indels    2;  Gaps    2;

Qy        344 NCTSDLCVHSASGPEDVALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTS 403
              |||||| ||:||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 NCTSDLXVHTASGPEDVALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTS 60

Qy        404 GFQPVSIKPSKADNPHLLTIQPDLS-TTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGG 462
              ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db         61 GFQPVSIKPSKADNPHLLTIQPDLSTTTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGG 120

Qy        463 RHTLHHSSPTSEAEEFVSRLSTQNYFRSLPRGTSNMTYGTFNFLGGRLMIPNTGISLLIP 522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RHTLHHSSPTSEAEEFVSRLSTQNYFRSLPRGTSNMTYGTFNFLGGRLMIPNTGISLLIP 180

Qy        523 PDAIPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEP 582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PDAIPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEP 240

Qy        583 SPDSWSLRLKKQSCEGSWEQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEAL 642
              ||||||| ||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        241 SPDSWSLALKKQSCEGSWE-DVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEAL 299

Qy        643 SVAAAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHF 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db        300 SVAAAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHL 359

Qy        703 KDSYHNLRLSIHDVPSSLWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLA 762
               |||||| || |||||||||||||||||||||||||||||||||||||||||||||||||
Db        360 XDSYHNLXLSXHDVPSSLWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLA 419

Qy        763 CKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIIS 822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        420 CKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIIS 479

Qy        823 SLDPPCRRGADWRTLAQKLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVA 882
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 SLDPPCRRGADWRTLAQKLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVA 539

Qy        883 GLGQPDAGLFTVSEAEC 899
              |       |   |||||
Db        540 GTPAGRWLLSQCSEAEC 556


RESULT 10
AAB50691
ID   AAB50691 standard; protein; 931 AA.
XX
AC   AAB50691;
XX
DT   19-MAR-2001  (first entry)
XX
DE   Human UNC5C protein SEQ ID NO:90.
XX
KW   Human; Caenorhabditis elegans; UNC-5; splice variant; nematode worm;
KW   protein-protein interaction; identification.
XX
OS   Homo sapiens.
XX
PN   WO200073328-A2.
XX
PD   07-DEC-2000.
XX
PF   02-JUN-2000; 2000WO-EP005108.
XX
PR   01-JUN-1999;   99GB-00012755.
XX
PA   (DEVG-) DEVGEN NV.
XX
PI   Van Criekinge W,  Roelens I,  Bogaert T,  Verwaerde P;
XX
DR   WPI; 2001-016508/02.
XX
PT   Three variants of human unc-5C cDNAs (unc-5Cb, unc-5Cc and unc-5C8) and a
PT   human unc-5HS1 cDNA, useful in yeast two hybrid experiments for
```

```
PT   identifying unknown human cDNAs which encode proteins that interact with
PT   the human unc-5C protein.
XX
PS   Disclosure; Page 224-227; 246pp; English.
XX
CC   The present invention describes 3 variants of human unc-5C cDNAs (unc-
CC   5Cb, unc-5Cc and unc-5C8) which correspond to alternatively spliced unc-
CC   5C transcripts, and a human unc-5HS1 cDNA which shares homology with the
CC   Rattus norvegicus unc-5HS1 cDNA. Also described are assays based on
CC   protein-protein-interactions between the unc-5 protein and a variety of
CC   different interacting proteins. The unc-5C variant cDNAs and unc-5HS1
CC   cDNA are useful in methods for identifying compounds which reduce or
CC   inhibit the lethal phenotype associated with the expression of the unc-5
CC   death domain in yeast. They are also useful in yeast two hybrid
CC   experiments for identifying unknown human cDNAs which encode proteins
CC   that interact with the human unc-5C protein. AAC90914 to AAC90971 and
CC   AAB50646 to AAB50693 represent sequences used in the exemplification of
CC   the present invention
XX
SQ   Sequence 931 AA;

  Query Match             55.8%;  Score 2673.5;  DB 4;  Length 931;
  Best Local Similarity   55.5%;  Pred. No. 1.1e-222;
  Matches  507;  Conservative  154;  Mismatches  221;  Indels   31;  Gaps   12;

Qy          9 PALLGIVLAAWLRGSGAQQS---ATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVC 65
Db         26 PAL--ALLSASGTGSAAQDDDFFHELPETFPSDPPEPLPHFLIEPEEAYIVKNKPVNLYC 83

Qy         66 KAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGEPTMEVRINVSRQQVEKVFGLEEYWCQ 125
Db         84 KASPATQIYFKCNSEWVHQKDHIVDERVDETSGLIVREVSIEISRQQVEELFGPEDYWCQ 143

Qy        126 CVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAEVEWLR 185
Db        144 CVAWSSAGTTKSRKAYVRIAYLRKTFEQEPLGKEVSLEQEVLLQCRPPEGIPVAEVEWLK 203

Qy        186 NEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNGGW 245
Db        204 NEDIIDPVEDRNFYITIDHNLIIKQARLSDTANYTCVAKNIVAKRKSTTATVIVYVNGGW 263

Qy        246 STWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDGSWSPW 305
Db        264 STWTEWSVCNSRCGRGYQKRTRTCTNPAPLNGGAFCEGQSV-QKIACTTLCPVDGRWTPW 322

Qy        306 SKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCVHSASGPEDVALYVG 365
Db        323 SKWSTCGTECTHWRRRECTAPAPKNGGKDCDGLVLQSKNCTDGLCMQTAPDSDDVALYVG 382

Qy        366 L-IAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHLLTIQ 424
Db        383 IVIAVIVCLAISVVVALFVYRKNHRDFESDIIDSSALNGGFQPVNIKAARQD---LLAVP 439

Qy        425 PDL-STTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHS----SPTSEAEEFV 479
Db        440 PDLTSAAAMYRGPVYALHD-VSDKIPMTNSPILDPLPNLKIKVYNTSGAVSPQDDLSEFT 498

Qy        480 SRLS---TQNYF---------RSLPRGT--SNMTYGTFNFLGGRLMIPNTGISLLIPPDA 525
Db        499 SKLSPQMTQSLLENEALSLKNQSLARQTDPSCTAFGSFNSLGGHLIVPNSGVSLLIPAGA 558

Qy        526 IPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPD 585
Db        559 IPQGRVYEMYVTVHRKETMRPPMDDSQTLLTPVVSCGPPGALLTRPVVLTMHHCADPNTE 618

Qy        586 SWSLRLKKQSCEGSWEQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA 645
Db        619 DWKILLKNQAAQGQWE-DVVVVGEENFTTPCYIKLDAEACHILTENLSTYALVGHSTTKA 677

Qy        646 AAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDS 705
Db        678 AAKRLKLAIFGPLCCSSLEYSIRVYCLDDTQDALKEILHLERQTGGQLLEEPKALHFKGS 737

Qy        706 YHNLRLSIHDVPSSLWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKL 765
Db        738 THNLRLSIHDIAHSLWKSKLLAKYQEIPFYHVWSGSQRNLHCTFTLERFSLNTVELVCKL 797

Qy        766 WVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLD 825
Db        798 CVRQVEGEGQIFQLNCTVSEEPTGIDLPLLDPANTITTVTGPSAFSIPLPIRQKLCSSLD 857

Qy        826 PPCRRGADWRTLAQKLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLG 885
Db        858 APQTRGHDWRMLAHKLNLDRYLNYFATKSSPTGVILDLWEAQNFPDGNLSMLAAVLEEMG 917

Qy        886 QPDAGLFTVSEAE 898
Db        918 RHETVVSLAAEGQ 930


RESULT 11
ADE63098
ID   ADE63098 standard; protein; 931 AA.
XX
AC   ADE63098;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Human Protein AAC67491, SEQ ID NO 9033.
XX
KW   Human; pain; neuronal tissue; gene therapy;
KW   spinal segmental nerve injury; chronic constriction injury; CCI;
KW   spared nerve injury; SNI; Chung.
XX
OS   Homo sapiens.
XX
PN   WO2003016475-A2.
XX
PD   27-FEB-2003.
XX
PF   14-AUG-2002; 2002WO-US025765.
XX
PR   14-AUG-2001; 2001US-0312147P.
PR   01-NOV-2001; 2001US-0346382P.
PR   26-NOV-2001; 2001US-0333347P.
XX
PA   (GEHO ) GEN HOSPITAL CORP.
PA   (FARB ) BAYER AG.
XX
PI   Woolf C,  D'urso D,  Befort K,  Costigan M;
XX
DR   WPI; 2003-268312/26.
DR   GENBANK; AAC67491.
XX
PT   New composition comprising two or more isolated polypeptides, useful for
PT   preparing a medicament for treating pain in an animal.
XX
PS   Claim 1; Page; 1017pp; English.
XX
CC   The invention discloses a composition comprising two or more isolated rat
CC   or human polynucleotides or a polynucleotide which represents a fragment,
CC   derivative or allelic variation of the nucleic acid sequence. Also
CC   claimed are a vector comprising the novel polynucleotide, a host cell
CC   comprising the vector, a method for identifying a nucleotide sequence
CC   which is differentially regulated in an animal subjected to pain and a
CC   kit to perform the method, an array, a method for identifying an agent
CC   that increases or decreases the expression of the polynucleotide sequence
CC   that is differentially expressed in neuronal tissue of a first animal
CC   subjected to pain, a method for identifying a compound which regulates
CC   the expression of a polynucleotide sequence which is differentially
CC   expressed in an animal subjected to pain, a method for identifying a
CC   compound that regulates the activity of one or more of the
CC   polynucleotides, a method for producing a pharmaceutical composition, a
CC   method for identifying a compound or small molecule that regulates the
CC   activity in an animal of one or more of the polypeptides given in the
CC   specification, a method for identifying a compound useful in treating
```

```
CC   pain and a pharmaceutical composition comprising the one or more
CC   polypeptides or their antibodies. The polynucleotide or the compound that
CC   modulates its activity is useful for preparing a medicament for treating
CC   pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC   injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC   therapy). The sequence presented is a human protein (shown in Table 2 of
CC   the specification) which is differentially expressed during pain. Note:
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic form directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 931 AA;

  Query Match             55.8%;  Score 2673.5;  DB 7;  Length 931;
  Best Local Similarity   55.5%;  Pred. No. 1.1e-222;
  Matches  507;  Conservative  154;  Mismatches  221;  Indels   31;  Gaps   12;

Qy          9 PALLGIVLAAWLRGSGAQQS---ATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVC 65
              |||   :|:|   || ||         :     |     |: |||||:|||: |||||||| |  |
Db         26 PAL--ALLSASGTGSAAQDDDFFHELPETFPSDPPEPLPHFLIEPEEAYIVKNKPVNLYC 83

Qy         66 KAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGEPTMEVRINVSRQQVEKVFGLEEYWCQ 125
              || |||||:|||| ||| | ||:::    | :||      || | :||||||::|| |:||||
Db         84 KASPATQIYFKCNSEWVHQKDHIVDERVDETSGLIVREVSIEISRQQVEELFGPEDYWCQ 143

Qy        126 CVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAEVEWLR 185
              ||||||:|||||:|||:||| ||| ||||||| |||||| ::| |||||||| |||||:
Db        144 CVAWSSAGTTKSRKAYVRIAYLRKTFEQEPLGKEVSLEQEVLLQCRPPEGIPVAEVEWLK 203

Qy        186 NEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNGGW 245
              |||::||  |  | |||  :|:|:::||||:||||||||||||||:|:| :| ||||||||
Db        204 NEDIIDPVEDRNFYITIDHNLIIKQARLSDTANYTCVAKNIVAKRKSTTATVIVYVNGGW 263

Qy        246 STWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDGSWSPW 305
              |||||||||:: |||:|||:|||:|||||||||||||||||:|     :    :    |  ||| |:||
Db        264 STWTEWSVCNSRCGRGYQKRTRTCTNPAPLNGGAFCEGQSV-QKIACTTLCPVDGRWTPW 322

Qy        306 SKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCVHSASGPEDVALYVG 365
              |||| || :||||| |||: |||:|||::| |   | ::|||  ||: :|     :||||||
Db        323 SKWSTCGTECTHWRRRECTAPAPKNGGKDCDGLVLQSKNCTDGLCMQTAPDSDDVALYVG 382

Qy        366 L-IAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHLLTIQ 424
              : ||| ||| : ::| : || :      :||: ||| |  |||||:|| :: |    || :
Db        383 IVIAVIVCLAISVVVALFVYRKNHRDFESDIIDSSALNGGFQPVNIKAARQD---LLAVP 439

Qy        425 PDL-STTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHS----SPTSEAEEFV 479
              ||| |      |:| :      |  |  | :||  :| ||      :  ::::       ||  :  ||
Db        440 PDLTSAAAMYRGPVYALHD-VSDKIPMTNSPILDPLPNLKIKVYNTSGAVSPQDDLSEFT 498

Qy        480 SRLS---TQNYF---------RSLPRGT--SNMTYGTFNFLGGRLMIPNTGISLLIPPDA 525
              |:||     ||:                :|| |  |     :|:|| ||| |::||:|:||||  |
Db        499 SKLSPQMTQSLLENEALSLKNQSLARQTDPSCTAFGSFNSLGGHLIVPNSGVSLLIPAGA 558

Qy        526 IPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPD 585
              ||:|::||:|:|:|:  | :| |:  ||||:|:||||||| ||||||:| | || :|: :
Db        559 IPQGRVYEMYVTVHRKETMRPPMDDSQTLLTPVVSCGPPGALLTRPVVLTMHHCADPNTE 618

Qy        586 SWSLRLKKQSCEGSWEQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA 645
               | : || |: :| || ||: :|||   :     | :|:| ||:: || |    :|||| : : |
Db        619 DWKILLKNQAAQGQWE-DVVVVGEENFTTPCYIKLDAEACHILTENLSTYALVGHSTTKA 677

Qy        646 AAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDS 705
              |||||| :| |: |:||||:|||||| || |||||:: ||:| ||||::||: |||| |
Db        678 AAKRLKLAIFGPLCCSSLEYSIRVYCLDDTQDALKEILHLERQTGGQLLEEPKALHFKGS 737

Qy        706 YHNLRLSIHDVPSSLWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKL 765
               |||||||||:  |||||||   |||||||:|:|:|| ||||||||| | :| :| |||
Db        738 THNLRLSIHDIAHSLWKSKLLAKYQEIPFYHVWSGSQRNLHCTFTLERFSLNTVELVCKL 797

Qy        766 WVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLD 825
               | ||||:|| |  :|   ::::     :|   |:      :    :  ||||| ||  ||||: ||||
Db        798 CVRQVEGEGQIFQLNCTVSEEPTGIDLPLLDPANTITTVTGPSAFSIPLPIRQKLCSSLD 857

Qy        826 PPCRRGADWRTLAQKLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLG 885
               |  || ||| || ||:|| :|::||:| ||| :||:||||::||:|||| ||| :  :|
Db        858 APQTRGHDWRMLAHKLNLDRYLNYFATKSSPTGVILDLWEAQNFPDGNLSMLAAVLEEMG 917

Qy        886 QPDAGLFTVSEAE 898
               : :  :   :| :
Db        918 RHETVVSLAAEGQ 930


RESULT 12
ABG11551
ID   ABG11551 standard; protein; 982 AA.
XX
AC   ABG11551;
XX
DT   18-FEB-2002  (first entry)
XX
DE   Novel human diagnostic protein #11542.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   N-PSDB; AAS75738.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 20; SEQ ID NO 41910; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC   amino acid sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 982 AA;

  Query Match             55.8%;  Score 2673.5;  DB 4;  Length 982;
  Best Local Similarity   55.5%;  Pred. No. 1.3e-222;
```

```
  Matches  507;  Conservative  154;  Mismatches  221;  Indels   31;  Gaps   12;

Qy          9 PALLGIVLAAWLRGSGAQQS---ATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVC 65
Db         77 PAL--ALLSASGTGSAAQDDDFFHELPETFPSDPPEPLPHFLIEPEEAYIVKNKPVNLYC 134

Qy         66 KAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGEPTMEVRINVSRQQVEKVFGLEEYWCQ 125
Db        135 KASPATQIYFKCNSEWVHQKDHIVDERVDETSGLIVREVSIEISRQQVEELFGPEDYWCQ 194

Qy        126 CVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAEVEWLR 185
Db        195 CVAWSSAGTTKSRKAYVRIAYLRKTFEQEPLGKEVSLEQEVLLQCRPPEGIPVAEVEWLK 254

Qy        186 NEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNGGW 245
Db        255 NEDIIDPVEDRNFYITIDHNLIIKQARLSDTANYTCVAKNIVAKRKSTTATVIVYVNGGW 314

Qy        246 STWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDGSWSPW 305
Db        315 STWTEWSVCNSRCGRGYQKRTRTCTNPAPLNGGAFCEGQSV-QKIACTTLCPVDGRWTPW 373

Qy        306 SKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCVHSASGPEDVALYVG 365
Db        374 SKWSTCGTECTHWRRRECTAPAPKNGGKDCDGLVLQSKNCTDGLCMQTAPDSDDVALYVG 433

Qy        366 L-IAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHLLTIQ 424
Db        434 IVIAVIVCLAISVVVALFVYRKNHRDFESDIIDSSALNGGFQPVNIKAARQD---LLAVP 490

Qy        425 PDL-STTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHS----SPTSEAEEFV 479
Db        491 PDLTSAAAMYRGPVYALHD-VSDKIPMTNSPILDPLPNLKIKVYNTSGAVSPQDDLSEFT 549

Qy        480 SRLS---TQNYF---------RSLPRGT--SNMTYGTFNFLGGRLMIPNTGISLLIPPDA 525
Db        550 SKLSPQMTQSLLENEALSLKNQSLARQTDPSCTAFGSFNSLGGHLIVPNSGVSLLIPAGA 609

Qy        526 IPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPD 585
Db        610 IPQGRVYEMYVTVHRKETMRPPMDDSQTLLTPVVSCGPPGALLTRPVVLTMHHCADPNTE 669

Qy        586 SWSLRLKKQSCEGSWEQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA 645
Db        670 DWKILLKNQAAQGQWE-DVVVVGEENFTTPCYIKLDAEACHILTENLSTYALVGHSTTKA 728

Qy        646 AAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDS 705
Db        729 AAKRLKLAIFGPLCCSSLEYSIRVYCLDDTQDALKEILHLERQTGGQLLEEPKALHFKGS 788

Qy        706 YHNLRLSIHDVPSSLWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKL 765
Db        789 THNLRLSIHDIAHSLWKSKLLAKYQEIPFYHVWSGSQRNLHCTFTLERFSLNTVELVCKL 848

Qy        766 WVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLD 825
Db        849 CVRQVEGEGQIFQLNCTVSEEPTGIDLPLLDPANTITTVTGPSAFSIPLPIRQKLCSSLD 908

Qy        826 PPCRRGADWRTLAQKLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLG 885
Db        909 APQTRGHDWRMLAHKLNLDRYLNYFATKSSPTGVILDLWEAQNFPDGNLSMLAAVLEEMG 968

Qy        886 QPDAGLFTVSEAE 898
Db        969 RHETVVSLAAEGQ 981
```

RESULT 13
ADE63096
ID   ADE63096 standard; protein; 945 AA.
XX
AC   ADE63096;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Rat Protein AAB57679, SEQ ID NO 9031.
XX
KW   Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW   chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
OS   Rattus norvegicus.
XX
PN   WO2003016475-A2.
XX
PD   27-FEB-2003.
XX
PF   14-AUG-2002; 2002WO-US025765.
XX
PR   14-AUG-2001; 2001US-0312147P.
PR   01-NOV-2001; 2001US-0346382P.
PR   26-NOV-2001; 2001US-0333347P.
XX
PA   (GEHO ) GEN HOSPITAL CORP.
PA   (FARB ) BAYER AG.
XX
PI   Woolf C,  D'urso D,  Befort K,  Costigan M;
XX
DR   WPI; 2003-268312/26.
DR   GENBANK; AAB57679.
XX
PT   New composition comprising two or more isolated polypeptides, useful for
PT   preparing a medicament for treating pain in an animal.
XX
PS   Claim 1; Page; 1017pp; English.
XX
CC   The invention discloses a composition comprising two or more isolated rat
CC   or human polynucleotides or a polynucleotide which represents a fragment,
CC   derivative or allelic variation of the nucleic acid sequence. Also
CC   claimed are a vector comprising the novel polynucleotide, a host cell
CC   comprising the vector, a method for identifying a nucleotide sequence
CC   which is differentially regulated in an animal subjected to pain and a
CC   kit to perform the method, an array, a method for identifying an agent
CC   that increases or decreases the expression of the polynucleotide sequence
CC   that is differentially expressed in neuronal tissue of a first animal
CC   subjected to pain, a method for identifying a compound which regulates
CC   the expression of a polynucleotide sequence which is differentially
CC   expressed in an animal subjected to pain, a method for identifying a
CC   compound that regulates the activity of one or more of the
CC   polynucleotides, a method for producing a pharmaceutical composition, a
CC   method for identifying a compound or small molecule that regulates the
CC   activity in an animal of one or more of the polypeptides given in the
CC   specification, a method for identifying a compound useful in treating
CC   pain and a pharmaceutical composition comprising the one or more
CC   polypeptides or their antibodies. The polynucleotide or the compound that
CC   modulates its activity is useful for preparing a medicament for treating
CC   pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC   injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC   therapy). The sequence presented is a rat protein (shown in Table 2 of
CC   the specification) which is differentially expressed during pain. Note:
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic form directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 945 AA;

```
  Query Match             52.2%;  Score 2497;  DB 7;  Length 945;
  Best Local Similarity   52.0%;  Pred. No. 2.6e-207;
  Matches  501;  Conservative  144;  Mismatches  236;  Indels   82;  Gaps   20;

Qy          1 MAVRPGLWPALLGIVLAAW-----LRG--SGAQQSATVANPVPGANPDLLPHFLVEPEDV 53
Db          1 MRARSGARGALLLALLLCWDPTPSLAGIDSGGQ---ALPDSFPSAPAEQLPHFLLEPEDA 57

Qy         54 YIVKNKPVLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGEPTMEVRINVSRQQV 113
Db         58 YIVKNKPVELHCRAFPATQIYFKCNGEWVSQKGHVTQESLDEATGLRIREVQIEVSRQQV 117
```

```
Qy        114 EKVFGLEEYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPP 173
              |::||||:|||||||||||||||::||||| |||||:|||||||  |:  ::| ||||
Db        118 EELFGLEDYWCQCVAWSSSGTTKSRRAYIRIAYLRKNFDQEPLAKEVPLDHEVLLQCRPP 177

Qy        174 EGIPPAEVEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSA 233
              ||:| |||||:|||::||: |  |  :| :|:|::||||:|||||||||||||||:|||
Db        178 EGVPVAEVEWLKNEDVIDPAQDTNFLLTIDHNLIIRQARLSDTANYTCVAKNIVAKRRST 237

Qy        234 SAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSS 293
               :| ||||||||||:| ||| ||  ||||||||:|:||||||||||||||     :|   :
Db        238 TATVIVYVNGGWSSWAEWSPCSNRCGRGWQKRTRTCTNPAPLNGGAFCEGQ-ACQKTACT 296

Qy        294 LLVSVDGSWSPWSKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCV-- 351
               :  |||:|: ||||||  :| |||||||  | |:|||  :| || ||::|||  |||
Db        297 TVCPVDGAWTEWSKWSACSTECAHWRSRECMAPPPQNGGRDCSGTLLDSKNCTDGLCVLN 356

Qy        352 -HSASGPE--------DVALYVGL-IAVAVCLVLLLLVLILVYCRKKEGLDSDVADSS-I 400
                : : |:         |||| || :|| |  | :|:  | ::|| |       |:|: |||
Db        357 QRTLNDPKSRPLEPSGDVALYAGLVVAVFVVLAVLMAVGVIVYRRNCRDFDTDITDSSAA 416

Qy        401 LTSGFQPVSIKPSKADNPHLL--TIQPDLSTTT-TYQGSLCPRQDGPSPKFQLTNGHLLS 457
              || || ||: | ::  || ||  :  |||: :    |:| :    ||  : |  :|| ||
Db        417 LTGGFHPVNFKTARPSNPQLLHPSAPPDLTASAGIYRGPVYALQDS-ADKIPMTNSPLLD 475

Qy        458 PL---------------GGG----------------------RHTLHHSSPTSEAEEFV 479
              ||                ||                       | ||  |
Db        476 PLPSLKIKVYDSSTIGSGAGLADGADLLGVLPPGTYPGDFSRDTHFLHLRS--------- 526

Qy        480 SRLSTQNYFRSLPRGTSNMTYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLH 539
               : | :|  :     |||  |:    |||   ||||| || ||:|||:|   |||:||  |::|| ::
Db        527 ASLGSQ-HLLGLPRDPSSSVSGTFGCLGGRLTIPGTGVSLLVPNGAIPQGKFYDLYLRIN 585

Qy        540 KPEDVRLPLA-GCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEG 598
              | |    |||: | ||:|||  |:|||  |:||  |||:| : ||  |    | :|| :| |
Db        586 KTEST-LPLSEGSQTVLSPSVTCGPTGLLLCRPVVLTVPHCAEVIAGDWIFQLKTQAHQG 644

Qy        599 SWEQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAAKRLKLLLFAPV 658
               || :|: | ||   :    ||||| :|::  :|||  :    ||:  |  :| |||:| :|||
Db        645 HWE-EVVTLDEETLNTPCYCQLEAKSCHILLDQLGTYVFTGESYSRSAVKRLQLAIFAPA 703

Qy        659 ACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPS 718
               ||||||::||||| ||  |||||::||: |||  |::||: | |||||||||||:||:|
Db        704 LCTSLEYSLRVYCLEDTPAALKEVLELERTLGGYLVEEPKTLLFKDSYHNLRLSLHDIPH 763

Qy        719 SLWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFS 778
               : |:||||  ||||||||:|||:|: |||||||||  | ::::  ||:  | ||||:|| |
Db        764 AHWRSKLLAKYQEIPFYHVWNGSQKALHCTFTLERHSLASTEFTCKVCVRQVEGEGQIFQ 823

Qy        779 INFNITKDTRFAELLALESEAGVPAL--VGPSAFKIPFLIRQKIISSLDPPCRRGADWRT 836
               ::  :  :|   ||  |  |   |   :||  |||||  |||||  :||| |   || |||
Db        824 LHTTLA-ETPAGSLDALCSAPGNAATTQLGPYAFKIPLSIRQKICNSLDAPNSRGNDWRL 882

Qy        837 LAQKLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSE 896
              ||||| :| :|::||:| ||| :||:|||||    :|:|: ||:|:  :|:  :  :    ::
Db        883 LAQKLSMDRYLNYFATKASPTGVILDLWEARQQDDGDLNSLASALEEMGKSEMLVAMTTD 942

Qy        897 AEC 899
               :|
Db        943 GDC 945
```

RESULT 14
ABG61795
ID   ABG61795 standard; protein; 946 AA.
XX
AC   ABG61795;
XX
DT   14-AUG-2002  (first entry)
XX
DE   Novel UNC5 receptor-like protein.
XX
KW   Cancer; metabolic disorder; skin disorder; infectious disease; anorexia;
KW   behavioral disorder; valve disease; endocrine disorder; heart disorder;
KW   blood disorder; anxiety disorder; brain disorder; inflammatory disorder;
KW   neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW   immune disorder; haematopoietic disorder; dyslipidaemia;
KW   metabolic disturbance; obesity; metabolic syndrome X; wasting disorder;
KW   gene therapy; transgenic animal; human.
XX
OS   Homo sapiens.
XX
PN   WO200229058-A2.
XX
PD   11-APR-2002.
XX
PF   05-OCT-2001; 2001WO-US031248.
XX
PR   05-OCT-2000; 2000US-0238323P.
PR   05-OCT-2000; 2000US-0238325P.
PR   06-OCT-2000; 2000US-0238372P.
PR   06-OCT-2000; 2000US-0238373P.
PR   06-OCT-2000; 2000US-0238379P.
PR   06-OCT-2000; 2000US-0238382P.
PR   06-OCT-2000; 2000US-0238383P.
PR   06-OCT-2000; 2000US-0238384P.
PR   06-OCT-2000; 2000US-0238397P.
PR   06-OCT-2000; 2000US-0238400P.
PR   06-OCT-2000; 2000US-0238401P.
PR   06-OCT-2000; 2000US-0238402P.
PR   14-MAR-2001; 2001US-0275892P.
PR   08-JUN-2001; 2001US-0296860P.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA, Taupier RJ, Burgess CE, Zerhusen BD, Mezes PS;
PI   Rastelli L, Malyankar UM, Grosse WM, Alsobrook JP, Lepley DM;
PI   Spytek KA, Li L, Edinger S, Gerlach V, Ellerman K, Macdougall J;
PI   Gunther E, Millet I, Stone D, Smithson G, Szekeres ES;
XX
DR   WPI; 2002-444103/47.
DR   N-PSDB; ABK92062.
XX
PT   Novel isolated polypeptide, designated NOVX, useful for treating or
PT   preventing cancer, diabetes, obesity, dyslipidemia, anorexia, and
PT   metabolic, neurodegenerative, immune and hematopoietic disorders.
XX
PS   Claim 1; Page 94; 316pp; English.
XX
CC   The invention describes an isolated polypeptide (I), useful in the
CC   manufacture of a medicament for treating a syndrome associated with a
CC   human disease. (I), the polynucleotide encoding it (II) and an antibody
CC   (III) to (I) are useful for treating or preventing cancer, metabolic
CC   disorders, skin disorders, infectious disease, anorexia, behavioral
CC   disorders, valve diseases, endocrine disorders, heart and blood
CC   disorders, anxiety disorders, brain disorders, inflammatory disorders,
CC   neurodegenerative disorder, Alzheimer's disease, Parkinson's disorder,
CC   immune disorders, haematopoietic disorders, and the various
CC   dyslipidaemias, metabolic disturbances associated with obesity, and the
CC   metabolic syndrome X and wasting disorders associated with chronic
CC   diseases. (I), (II) or (III) are useful in screening assays, detection
CC   assays (e.g., chromosomal mapping, tissue typing, forensic biology),
CC   predictive medicine (e.g., diagnostic assays and prognostic assays), and
CC   in methods of treatment (e.g., therapeutic and prophylactic). (I) is
CC   useful as immunogen to produce antibodies immunospecific for (I), to
CC   screen for potential agonist and antagonist compounds, and as bait
CC   protein in a two-hybrid or three-hybrid assay. (II) is useful in gene
CC   therapy, to express (I), to detect (I) mRNA or a genetic lesion in a (I)
CC   gene, and to modulate activity of (I). A cell containing a vector
CC   exressing (I) is useful for producing non-human transgenic animals. This
CC   is the amino acid sequence of a novel human polypeptide described in the
CC   invention
XX
SQ   Sequence 946 AA;

Query Match          52.0%;  Score 2490.5;  DB 5;  Length 946;

```
Best Local Similarity   52.1%;  Pred. No. 9.7e-207;
Matches  496;  Conservative  150;  Mismatches  247;  Indels   59;  Gaps   18;

Qy          1 MAVRPGLWPALLGIVLAAW---LRGSGAQQSATV-ANPVPGANPDLLPHFLVEPEDVYIV 56
Db          1 MGARSGARGALLLALLLCWDPRLSQAGRKRSGEVLPDSFPSAPAEPLPYFLQEPQDAYIV 60

Qy         57 KNKPVLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSS-----GEPTMEVRINVSRQ 111
Db         61 KNKPVELRCRAFPATQIYFKCNGEWVSQNDHVTQEGLDEATLGARGGLRVREVQIEVSRQ 120

Qy        112 QVEKVFGLEEYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCR 171
Db        121 QVEELFGLEDYWCQCVAWSSAGTTKSRRAYVRIACLRKNFDQEPLGKEVPLDHEVLLQCR 180

Qy        172 PPEGIPPAEVEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRR 231
Db        181 PPEGVPVAEVEWLKNEDVIDPTQDTNFLLTIDHNLIIRQARLSDTANYTCVAKNIVAKRR 240

Qy        232 SASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTV 291
Db        241 STTATVIVYVNGGWSSWAEWSPCSNRCGRGWQKRTRTCTNPAPLNGGAFCEGQ-AFQKTA 299

Qy        292 SSLLVSVDGSWSPWSKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCV 351
Db        300 CTTICPVDGAWTEWSKWSACSTECAHWRSRECMAPPPQNGGRDCSGTLLDSKNCTDGLCM 359

Qy        352 HS--------ASGPEDVALYVGL-IAVAVCLVLLLLVLILVYCRKKEGLDSDVADSS-IL 401
Db        360 QSEPVPAVLEASG--DAALYAGLVVAIFVVVAILMAVGVVVYRRNCRDFDTDITDSSAAL 417

Qy        402 TSGFQPVSIKPSKADNPHLL--TIQPDLSTTT-TYQGSLCPRQDGPSPKFQLTNGHLLSP 458
Db        418 TGGFHPVNFKTARPSNPQLLHPSVPPDLTASAGIYRGPVYALQDS-TDKIPMTNSPLLDP 476

Qy        459 LGGGRHTLHHSSPT------SEAEEFVSRLSTQNY------------FRS---------- 490
Db        477 LPSLKVKVYSSSTTGSGPGLADGADLLGVLPPGTYPSDFARDTHFLHLRSASLGSQQLLG 536

Qy        491 LPRGTSNMTYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLHKPEDVRLPLA- 549
Db        537 LPRDPGSSVSGTFGCLGGRLSIPGTGVSLLVPNGAIPQGKFYEMYLLINKAEST-LPLSE 595

Qy        550 GCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGSWEQDVLHLGE 609
Db        596 GTQTVLSPSVTCGPTGLLLCRPVILTMPHCAEVSARDWIFQLKTQAHQGHWEQEVVTLDE 655

Qy        610 EAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAAKRLKLLLFAPVACTSLEYNIRV 669
Db        656 ETLNTPCYCQLEPRACHILLDQLGTYVFTGESYSRSAVKRLQLAVFAPALCTSLEYSLRV 715

Qy        670 YCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWKSKLLVSY 729
Db        716 YCLEDTPVALKEVLELERTLGGYLVEEPKPLMFKDSYHNLRLSLHDLPHAHWRSKLLAKY 775

Qy        730 QEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRF 789
Db        776 QEIPFYHIWSGSQKALHCTFTLERHSLASTELTCKICVRQVEGEGQIFQLHTTLA-ETPA 834

Qy        790 AELLALESEAG--VPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLDSHL 847
Db        835 GSLDTLCSAPGSTVTTQLGPYAFKIPLSIRQKICNSLDAPNSRGNDWRMLAQKLSMDRYL 894

Qy        848 SFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 899
Db        895 NYFATKASPTGVILDLWEALQQDDGDLNSLASALEEMGKSEMLVAVATDGDC 946
```

```
RESULT 15
AAW78900
ID   AAW78900 standard; protein; 943 AA.
XX
AC   AAW78900;
XX
DT   25-MAR-2003  (revised)
DT   21-DEC-1998  (first entry)
XX
DE   Rat UNC-5 homologue UNC5H-2.
XX
KW   UNC-5; UNC5H-2; rat; netrin receptor; cell migration; axon guidance;
KW   diagnosis; therapy.
XX
OS   Rattus sp.
XX
FH   Key             Location/Qualifiers
FT   Peptide         148. .161
FT                   /note= "peptide used to raise rabbit polyclonal antisera"
FT   Misc-difference 753
FT                   /note= "encoded by CG"
FT   Peptide         909. .924
FT                   /note= "peptide used to raise rabbit polyclonal antisera"
XX
PN   WO9837085-A1.
XX
PD   27-AUG-1998.
XX
PF   19-FEB-1998;   98WO-US003143.
XX
PR   19-FEB-1997;   97US-00808982.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Tessier-Lavigne M,  Leonardo ED,  Hinck L,  Masu M,  Keinomasu K;
XX
DR   WPI; 1998-495364/42.
DR   N-PSDB; AAV52942.
XX
PT   Netrin-binding, vertebrate proteins - useful for diagnosis, therapy and
PT   the biopharmaceutical industry.
XX
PS   Claim 1; Page 24-26; 32pp; English.
XX
CC   UNC5H-1 and UNC5H-2 (see AAW78900) are rat homologues of Caenorhabditis
CC   elegans UNC-5 protein. Their amino acid sequences were deduced from
CC   isolated unc5h cDNA clones (see AAV52940 and AAV52942) isolated from an
CC   E18 brain cDNA library. The predicted proteins show similarity with UNC-
CC   5, possess 2 predicted Ig-like domains and 2 predicted thrombospondin
CC   type-1 repeats, a predicted membrane spanning region, and a large
CC   intracellular domain. They are predicted to be involved in cell migration
CC   and axon guidance, and are characterised as receptor proteins for
CC   netrins. Human UNC5H-1 (see AAW78899) and UNC5H-2 (see AAW78901) proteins
CC   are also claimed. Vertebrate UNC-5 proteins may be produced recombinantly
CC   from transfected host cells. The invention also provides unc-5
CC   hybridisation probes and primers, vertebrate UNC-5-specific binding
CC   agents such as specific antibodies, and methods of making and using the
CC   subject compositions in diagnosis (e.g. genetic hybridisation screens for
CC   vertebrate unc-5 transcripts), therapy (e.g. gene therapy to modulate
CC   vertebrate unc-5 gene expression) and in the biopharmaceutical industry
CC   (e.g. as immunogens, reagents for modulating cell guidance, reagents for
CC   screening chemical libraries for lead pharmacological agents, etc.).
CC   (Updated on 25-MAR-2003 to correct PI field.)
XX
SQ   Sequence 943 AA;

  Query Match             52.0%;  Score 2490;  DB 2;  Length 943;
  Best Local Similarity   52.3%;  Pred. No. 1.1e-206;
  Matches  496;  Conservative  144;  Mismatches  226;  Indels   82;  Gaps   19;

Qy          9 PALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAV 68
Db         21 PSLAGI-------DSGAQ---GLPDSFPSAPAEQLPHFLLEPEDAYIVKNKPVELHCRAF 70

Qy         69 PATQIFFKCNGEWVRQVDHVIERSTDGSSGEPTMEVRINVSRQQVEKVFGLEEYWCQCVA 128
Db         71 PATQIYFKCNGEWVSQKGHVTQESLDEATGLRIREVQIEVSRQQVEELFGLEDYWCQCVA 130
```

```
Qy   129 WSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAEVEWLRNED 188
Db   131 WSSSGTTKSRRAYIRIAYLRKNFDQEPLAKEVPLDHEVLLQCRPPEGVPVAEVEWLKNED 190

Qy   189 LVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNGGWSTW 248
Db   191 VIDPAQDTNFLLTIDHNLIIRQARLSDTANYTCVAKNIVAKRRSTTATVIVYVNGGWSSW 250

Qy   249 TEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDGSWSPWSKW 308
Db   251 AEWSPCSNRCGRGWQKRTRTCTNPAPLNGGAFCEGQ-ACQKTACTTVCPVDGAWTEWSKW 309

Qy   309 SACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCV---HSASGPE------- 358
Db   310 SACSTECAHWRSRECMAPPPQNGGRDCSGTLLDSKNCTDGLCVLNQRTLNDPKSRPLEPS 369

Qy   359 -DVALYVGL-IAVAVCLVLLLLVLILVYCRKKEGLDSDVADSS-ILTSGFQPVSIKPSKA 415
Db   370 GDVALYAGLVVAVFVVLAVLMAVGVIVYRRNCRDFDTDITDSSAALTGGFHPVNFKTARP 429

Qy   416 DNPHLL--TIQPDLSTTT-TYQGSLCPRQDGPSPKFQLTNGHLLSPL------------- 459
Db   430 SNPQLLHPSAPPDLTASAGIYRGPVYALQDS-ADKIPMTNSPLLDPLPSLKIKVYDSSTI 488

Qy   460 --GGG----------------------RHTLHHSSPTSEAEEFVSRLSTQNYFRSLPRG 494
Db   489 GSGAGLADGADLLGVLPPGTYPGDFSRDTHFLHLRS---------ASLGSQ-HLLGLPRD 538

Qy   495 TSNMTYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLHKPEDVRLPLA-GCQT 553
Db   539 PSSSVSGTFGCLGGRLTIPGTGVSLLVPNGAIPQGKFYDLYLRINKTEST-LPLSEGSQT 597

Qy   554 LLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGSWEQDVLHLGEEAPS 613
Db   598 VLSPSVTCGPTGLLLCRPVVLTVPHCAEVIAGDWIFQLKTQAHQGHWE-EVVTLDEETLN 656

Qy   614 HLYYCQLEASACYVFTEQLGRFALVGEALSVAAAKRLKLLLFAPVACTSLEYNIRVYCLH 673
Db   657 TPCYCQLEAKSCHILLDQLGTYVFTGESYSRSAVKRLQLAIFAPALCTSLEYSLRVYCLE 716

Qy   674 DTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWKSKLLVSYQEIP 733
Db   717 DTPAALKEVLELERTLGGYLVEEPKTLLFKDSYHNLRLSLHDIPHAHWRSKLLAKYQEIP 776

Qy   734 FYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELL 793
Db   777 FYHVWNGSQKALHCTFTLERHSLASTEFTCKVCVRQVEGEGQIFQLHTTLA-ETPAGSLD 835

Qy   794 ALESEAGVPAL--VGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLDSHLSFFA 851
Db   836 ALCSAPGNAATTQLGPYAFKIPLSIRQKICNSLDAPNSRGNDWRLLAQKLSMDRYLNYFA 895

Qy   852 SKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 899
Db   896 TKASPTGVILDLWEARQQDDGDLNSLASALEEMGKSEMLVAMTTDGDC 943
```

Search completed: October  6, 2004, 18:17:43
Job time : 103 secs

```
                  GenCore version 5.1.6
           Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:         October  6, 2004, 18:20:44 ; Search time 152 Seconds
                                             (without alignments)
                                             3282.242 Million cell updates/sec

Title:          US-09-970-944-2
Perfect score:  4787
Sequence:       1 MAVRPGLWPALLGIVLAAWL..........AVAGLGQPDAGLFTVSEAEC 899

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters:      1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09970944/runat_05102004_112011_1655/app_query.fasta_1.1095
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09970944_@CGN_1_1_75_@runat_05102004_112011_1655 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      Issued_Patents_NA:*
                1:  /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
                2:  /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
                3:  /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
                4:  /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
                5:  /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
                6:  /cgn2_6/ptodata/2/ina/backfiles1.seq:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.
```

SUMMARIES

| | Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|---|
| | 1 | 4545.5 | 95.0 | 3014 | 2 | US-08-808-982-1 | Sequence 1, Appli |
| | 2 | 4545.5 | 95.0 | 3014 | 3 | US-09-306-902A-1 | Sequence 1, Appli |
| | 3 | 2694 | 56.3 | 1787 | 2 | US-08-808-982-2 | Sequence 2, Appli |
| | 4 | 2694 | 56.3 | 1787 | 3 | US-09-306-902A-2 | Sequence 2, Appli |
| | 5 | 2463 | 51.5 | 2831 | 2 | US-08-808-982-3 | Sequence 3, Appli |
| | 6 | 2463 | 51.5 | 2831 | 3 | US-09-306-902A-3 | Sequence 3, Appli |
| | 7 | 2039 | 42.6 | 1282 | 4 | US-09-833-381-1806 | Sequence 1806, Ap |
| | 8 | 1123 | 23.5 | 1605 | 4 | US-09-833-381-1807 | Sequence 1807, Ap |
| c | 9 | 600.5 | 12.5 | 771 | 1 | US-08-253-155A-17 | Sequence 17, Appl |
| | 10 | 267.5 | 5.6 | 5749 | 4 | US-09-262-537-48 | Sequence 48, Appl |
| | 11 | 253 | 5.3 | 305 | 2 | US-08-808-982-4 | Sequence 4, Appli |
| | 12 | 253 | 5.3 | 305 | 3 | US-09-306-902A-4 | Sequence 4, Appli |
| | 13 | 250 | 5.2 | 1326 | 3 | US-08-985-526-4 | Sequence 4, Appli |
| | 14 | 250 | 5.2 | 7231 | 4 | US-09-919-172-64 | Sequence 64, Appl |
| | 15 | 245.5 | 5.1 | 657 | 3 | US-08-985-526-2 | Sequence 2, Appli |
| | 16 | 211.5 | 4.4 | 3874 | 4 | US-09-877-730-31 | Sequence 31, Appl |
| | 17 | 204.5 | 4.3 | 3991 | 4 | US-08-506-296B-3 | Sequence 3, Appli |
| | 18 | 204 | 4.3 | 3210 | 4 | US-09-877-730-1 | Sequence 1, Appli |
| | 19 | 204 | 4.3 | 3453 | 4 | US-09-877-730-7 | Sequence 7, Appli |
| | 20 | 192.5 | 4.0 | 2715 | 4 | US-09-877-730-5 | Sequence 5, Appli |
| | 21 | 192.5 | 4.0 | 2958 | 4 | US-09-877-730-9 | Sequence 9, Appli |
| | 22 | 186 | 3.9 | 1143 | 4 | US-09-877-730-3 | Sequence 3, Appli |
| | 23 | 184 | 3.8 | 7702 | 4 | US-09-023-655-1336 | Sequence 1336, Ap |
| | 24 | 178.5 | 3.7 | 3675 | 4 | US-09-930-872-3 | Sequence 3, Appli |
| | 25 | 178.5 | 3.7 | 4042 | 4 | US-09-930-872-5 | Sequence 5, Appli |
| c | 26 | 175.5 | 3.7 | 8312 | 4 | US-09-620-312D-1048 | Sequence 1048, Ap |
| | 27 | 173.5 | 3.6 | 6814 | 4 | US-09-484-970B-66 | Sequence 66, Appl |
| c | 28 | 170.5 | 3.6 | 3126 | 4 | US-09-392-184-7 | Sequence 7, Appli |
| | 29 | 170.5 | 3.6 | 5690 | 2 | US-08-447-464-2 | Sequence 2, Appli |
| | 30 | 170.5 | 3.6 | 5690 | 2 | US-08-716-679-2 | Sequence 2, Appli |
| | 31 | 168 | 3.5 | 2114 | 4 | US-09-130-491-7 | Sequence 7, Appli |
| | 32 | 168 | 3.5 | 4548 | 4 | US-09-571-479C-5 | Sequence 5, Appli |
| | 33 | 167.5 | 3.5 | 3638 | 4 | US-09-369-364A-8 | Sequence 8, Appli |
| | 34 | 166 | 3.5 | 4192 | 4 | US-09-122-126B-1 | Sequence 1, Appli |
| | 35 | 166 | 3.5 | 4192 | 4 | US-09-634-286A-1 | Sequence 1, Appli |
| | 36 | 165 | 3.4 | 4843 | 3 | US-08-986-485-1 | Sequence 1, Appli |
| | 37 | 164.5 | 3.4 | 3218 | 4 | US-09-369-364A-6 | Sequence 6, Appli |
| | 38 | 163 | 3.4 | 4078 | 4 | US-09-016-434-1132 | Sequence 1132, Ap |
| | 39 | 163 | 3.4 | 6000 | 1 | US-08-348-006B-6 | Sequence 6, Appli |
| | 40 | 163 | 3.4 | 6000 | 2 | US-08-800-825A-6 | Sequence 6, Appli |
| | 41 | 163 | 3.4 | 6000 | 3 | US-09-158-657-6 | Sequence 6, Appli |
| | 42 | 163 | 3.4 | 6000 | 5 | PCT-US94-10166-6 | Sequence 6, Appli |
| c | 43 | 162.5 | 3.4 | 5357 | 4 | US-09-392-184-5 | Sequence 5, Appli |
| | 44 | 160.5 | 3.4 | 3889 | 4 | US-09-568-559-1 | Sequence 1, Appli |
| | 45 | 160.5 | 3.4 | 4676 | 4 | US-09-130-491-1 | Sequence 1, Appli |

ALIGNMENTS

```
RESULT 1
US-08-808-982-1
; Sequence 1, Application US/08808982
; Patent No. 5939271
;  GENERAL INFORMATION:
;    APPLICANT:  Tessier-Lavigne, Marc
;    APPLICANT:  Leonardo, E. David
;    APPLICANT:  Hink, Lindsay
;    APPLICANT:  Masu, Masayuki
;    APPLICANT:  Kazuko, Keino-Masu
;    TITLE OF INVENTION:  Netrin Receptors
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SCIENCE & TECHNOLOGY LAW GROUP
;      STREET:  268 BUSH STREET, SUITE 3200
;      CITY:  SAN FRANCISCO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/808,982
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  OSMAN, RICHARD A
;      REGISTRATION NUMBER:  36,627
;      REFERENCE/DOCKET NUMBER:  UC96-217
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 343-4341
;      TELEFAX:  (415) 343-4342
;  INFORMATION FOR SEQ ID NO:  1:
```

```
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3014 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-808-982-1

Alignment Scores:
Pred. No.:                0             Length:          3014
Score:                    4545.50       Matches:         852
Percent Similarity:       96.78%        Conservative:    19
Best Local Similarity:    94.67%        Mismatches:      26
Query Match:              94.96%        Indels:          3
DB:                       2             Gaps:            3

US-09-970-944-2 (1-899) x US-08-808-982-1 (1-3014)

Qy      1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeu 20
Db      1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGTGCTCCTGGGCATAGTCCTCGCCGCCTGGCTT 60

Qy     21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40
Db     61 CGTGGTTCGGGTGCCCAGCAGAGTGCCACGGTGGCCAATCCAGTGCCCGGTGCCAACCCC 120

Qy     41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro 60
Db    121 GACCTGCTGCCCCACTTCCTGGTAGAGCCTGAGGACGTGTACATTGTCAAGAACAAGCCG 180

Qy     61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80
Db    181 GTGTTGTTGGTGTGCAAGGCTGTGCCTGCCACCCAGATCTTCTTCAAGTGCAATGGGGAA 240

Qy     81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyPro 100
Db    241 TGGGTCCGCCAGGTCGATCACGTAATTGAACGCAGCACCGACAGCAGCAGCGGATTGCCA 300

Qy    101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120
Db    301 ACCATGGAGGTCCGTATCAACGTATCGAGGCAGCAGGTAGAGAAAGTGTTTGGGCTGGAG 360

Qy    121 GluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAla 140
Db    361 GAATACTGGTGCCAGTGTGTGGCATGGAGCTCCTCGGGTACCACCAAAAGTCAGAAGGCC 420

Qy    141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160
Db    421 TACATCCGGATTGCCTATTTGCGCAAGAACTTTGAGCAGGAGCCACTGGCCAAGGAAGTG 480

Qy    161 SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu 180
Db    481 TCACTGGAGCAAGGCATTGTACTACCTTGTCGCCCCCCAGAAGGAATCCCCCCAGCTGAG 540

Qy    181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200
Db    541 GTGGAGTGGCTTCGAAATGAGGACCTCGTGGACCCCTCCCTCGATCCCAATGTGTACATC 600

Qy    201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220
Db    601 ACGCGGGAGCACAGCCTAGTCGTGCGTCAGGCCCGCCTGGCCGACACGGCCAACTACACC 660

Qy    221 CysValAlaLysAsnIleValAlaArgArgArgSerAlaSerAlaAlaValIleValTyr 240
Db    661 TGTGTGGCCAAGAACATCGTAGCCCGTCGCCGAAGCACCTCTGCAGCGGTCATTGTTTAT 720

Qy    241 ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg 260
Db    721 GTGAACGGTGGGTGGTCGACGTGGACTGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGT 780

Qy    261 GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe 280
Db    781 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCACCTCTCAACGGGGGCGCCTTC 840

Qy    281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300
Db    841 TGTGAGGGGCAGAATGTC---CAGAAAACAGCCTGCGCCACTCTGTGCCCAGTGGATGGG 897

Qy    301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320
Db    898 AGCTGGAGTTCGTGGAGTAAGTGGTCAGCCTGTGGGCTTGACTGCACCCACTGGCGGAGC 957

Qy    321 ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeu 340
Db    958 CGCGAGTGCTCTGACCCAGCACCCCGCAATGGAGGTGAGGAGTGTCGGGGTGCTGACCTG 1017

Qy    341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360
Db   1018 GACACCCGCAACTGTACCAGTGACCTCTGCCTGCACACCGCTTCTTGCCCCGAGGACGTG 1077

Qy    361 AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuLeuValLeu 380
Db   1078 GCTCTCTACATCGGCCTTGTCGCTGTGGCTGTGTGCCTCTTCTTGCTGTTGCTGGCCCTT 1137

Qy    381 IleLeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIle 400
Db   1138 GGACTCATTTACTGTCGCAAGAAGGAAGGGCTGGACTCCGATGTGGCCGACTCGTCCATC 1197

Qy    401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420
Db   1198 CTCACCTCGGGCTTCCAGCCTGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCACCTG 1257

Qy    421 LeuThrIleGlnProAspLeuSer---ThrThrThrThrTyrGlnGlySerLeuCysPro 439
Db   1258 CTCACCATCCAGCCAGACCTCAGCACCACCACTACCACCTACCAGGGCAGTCTATGTTCG 1317

Qy    440 ArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu 459
Db   1318 AGGCAGGATGGACCCAGCCCCAAGTTCCAGCTCTCTAATGGTCACCTGCTCAGCCCACTG 1377

Qy    460 GlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheVal 479
Db   1378 GGGAGTGGCCGCCATACGTTGCACCACAGCTCACCCACCTCTGAGGCTGAGGACTTCGTC 1437

Qy    480 SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr 499
Db   1438 TCCCGCCTCTCCACCCAAAACTACTTTCGTTCCCTGCCCCGCGGCACCAGCAACATGGCC 1497

Qy    500 TyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeu 519
Db   1498 TACGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACGGGGATCAGCCTC 1557

Qy    520 LeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHis 539
Db   1558 CTCATACCCCCGGATGCCATCCCCCGAGGAAAGATCTACGAGATCTACCTCACACTGCAC 1617

Qy    540 LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal 559
Db   1618 AAGCCAGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCAGTCGTT 1677

Qy    560 SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys 579
Db   1678 AGCTGTGGGCCCCCAGGAGTCCTGCTCACCCGGCCAGTCATCCTTGCAATGGACCACTGT 1737

Qy    580 GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySer 599
Db   1738 GGAGAGCCCAGCCCTGACAGCTGGAGTCTGCGCCTCAAAAAGCAGTCCTGCGAGGGCAGT 1797

Qy    600 TrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGln 619
Db   1798 TGGGAG---GATGTGCTGCACCTTGGTGAGGAGTCACCTTCCCACCTCTACTACTGCCAG 1854

Qy    620 LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly 639
Db   1855 CTGGAGGCCGGGGCCTGCTATGTCTTCACGGAGCAGCTGGGCCGCTTTGCCCTGGTAGGA 1914
```

```
Qy      640 GluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAla 659
Db     1915 GAGGCCCTCAGCGTGGCTGCCACCAAGCGCCTCAGGCTCCTTCTGTTTGCTCCCGTGGCC 1974

Qy      660 CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu 679
Db     1975 TGTACGTCCCTTGAGTACAACATCCGAGTGTACTGCCTACACGACACCCACGACGCTCTC 2034

Qy      680 LysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgVal 699
Db     2035 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTAGGTGGACAGCTGATCCAGGAGCCTCGCGTC 2094

Qy      700 LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719
Db     2095 CTGCACTTCAAAGACAGTTACCACAACCTACGTCTCTCCATCCACGACGTGCCCAGCTCC 2154

Qy      720 LeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn 739
Db     2155 CTGTGGAAGAGCAAGCTACTTGTCAGCTACCAGGAGATCCCTTTTTACCACATCTGGAAC 2214

Qy      740 GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759
Db     2215 GGCACCCAGCAGTATCTGCACTGCACCTTCACCCTGGAGCGCATCAACGCCAGCACCAGC 2274

Qy      760 AspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIle 779
Db     2275 GACCTGGCCTGCAAGGTGTGGGTGTGGCAGGTGGAGGGAGATGGGCAGAGCTTCAACATC 2334

Qy      780 AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAla 799
Db     2335 AACTTCAACATCACTAAGGACACAAGGTTTGCTGAATTGTTGGCTCTGGAGAGTGAAGGG 2394

Qy      800 GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
Db     2395 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAAAAG 2454

Qy      820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839
Db     2455 ATCATCGCCAGTCTGGACCCACCCTGCAGCCGGGGCGCCGACTGGAGAACTCTAGCCCAG 2514

Qy      840 LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859
Db     2515 AAACTTCACCTGGACAGCCATCTTAGCTTCTTTGCCTCCAAGCCCAGCCCTACAGCCATG 2574

Qy      860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879
Db     2575 ATCCTCAACCTATGGGAGGCACGGCACTTCCCCAACGGCAACCTCGGCCAGCTGGCAGCA 2634

Qy      880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
Db     2635 GCTGTGGCCGGACTGGGCCAACCAGATGCTGGCCTCTTCACGGTGTCGGAGGCCGAGTGT 2694

RESULT 2
US-09-306-902A-1
; Sequence 1, Application US/09306902A
; Patent No. 6277585
;  GENERAL INFORMATION:
;       APPLICANT: Tessier-Lavigne, Marc
;                  Leonardo, E. David
;                  Hink, Lindsay
;                  Masu, Masayuki
;                  Kazuko, Keino-Masu
;       TITLE OF INVENTION: Netrin Receptors
;       NUMBER OF SEQUENCES: 9
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
;            STREET: 268 BUSH STREET, SUITE 3200
;            CITY: SAN FRANCISCO
;            STATE: CALIFORNIA
;            COUNTRY: USA
;            ZIP: 94104
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/306,902A
;            FILING DATE: 07-May-1999
;            CLASSIFICATION: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: OSMAN, RICHARD A
;            REGISTRATION NUMBER: 36,627
;            REFERENCE/DOCKET NUMBER: UC96-217
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (415) 343-4341
;            TELEFAX: (415) 343-4342
;  INFORMATION FOR SEQ ID NO: 1:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 3014 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: double
;            TOPOLOGY: linear
;       MOLECULE TYPE: cDNA
;       SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-306-902A-1

Alignment Scores:
Pred. No.:              0               Length:         3014
Score:                  4545.50         Matches:        852
Percent Similarity:     96.78%          Conservative:   19
Best Local Similarity:  94.67%          Mismatches:     26
Query Match:            94.96%          Indels:         3
DB:                     3               Gaps:           3

US-09-970-944-2 (1-899) x US-09-306-902A-1 (1-3014)

Qy        1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeu 20
Db        1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGTGCTCCTGGGCATAGTCCTCGCCGCCTGGCTT 60

Qy       21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40
Db       61 CGTGGTTCGGGTGCCCAGCAGAGTGCCACGGTGGCCAATCCAGTGCCCGGTGCCAACCCC 120

Qy       41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro 60
Db      121 GACCTGCTGCCCCACTTCCTGGTAGAGCCTGAGGACGTGTACATTGTCAAGAACAAGCCG 180

Qy       61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80
Db      181 GTGTTGTTGGTGTGCAAGGCTGTGCCTGCCACCCAGATCTTCTTCAAGTGCAATGGGGAA 240

Qy       81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100
Db      241 TGGGTCCGCCAGGTCGATCACGTAATTGAACGCAGCACCGACAGCAGCAGCGGATTGCCA 300

Qy      101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120
Db      301 ACCATGGAGGTCCGTATCAACGTATCGAGGCAGCAGGTAGAGAAAGTGTTTGGGCTGGAG 360

Qy      121 GluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAla 140
Db      361 GAATACTGGTGCCAGTGTGTGGCATGGAGCTCCTCGGGTACCACCAAAAGTCAGAAGGCC 420

Qy      141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160
Db      421 TACATCCGGATTGCCTATTTGCGCAAGAACTTTGAGCAGGAGCCACTGGCCAAGGAAGTG 480

Qy      161 SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu 180
Db      481 TCACTGGAGCAAGGCATTGTACTACCTTGTCGCCCCCCAGAAGGAATCCCCCCAGCTGAG 540

Qy      181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200
Db      541 GTGGAGTGGCTTCGAAATGAGGACCTCGTGGACCCCTCCCTCGATCCCAATGTGTACATC 600
```

```
Qy     201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 ACGCGGGAGCACAGCCTAGTCGTGCGTCAGGCCCGCCTGGCCGACACGGCCAACTACACC 660

Qy     221 CysValAlaLysAsnIleValAlaArgArgArgSerAlaSerAlaAlaValIleValTyr 240
           ||||||||||||||||||||||||||||||||||||   |||||||||||||||||||||
Db     661 TGTGTGGCCAAGAACATCGTAGCCCGTCGCCGAAGCACCTCTGCAGCGGTCATTGTTTAT 720

Qy     241 ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg 260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 GTGAACGGTGGGTGGTCGACGTGGACTGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGT 780

Qy     261 GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe 280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCACCTCTCAACGGGGGCGCCTTC 840

Qy     281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300
           ||||||||||||||||||      :::|||      :::   |||      |||||||||
Db     841 TGTGAGGGGCAGAATGTC---CAGAAAACAGCCTGCGCCACTCTGTGCCCAGTGGATGGG 897

Qy     301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320
           |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db     898 AGCTGGAGTTCGTGGAGTAAGTGGTCAGCCTGTGGGCTTGACTGCACCCACTGGCGGAGC 957

Qy     321 ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeu 340
           |||||||||||||||||||||||||||||||||||||||||||||:::|||   ||||||
Db     958 CGCGAGTGCTCTGACCCAGCACCCCGCAATGGAGGTGAGGAGTGTCGGGGTGCTGACCTG 1017

Qy     341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360
           ||||||||||||||||||||||||||||||:::|||:::||||||   ||||||||||||
Db    1018 GACACCCGCAACTGTACCAGTGACCTCTGCCTGCACACCGCTTCTTGCCCCGAGGACGTG 1077

Qy     361 AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuLeuValLeu 380
           |||||||||:::||||||:::||||||||||||||||||   ||||||||||||   |||
Db    1078 GCTCTCTACATCGGCCTTGTCGCTGTGGCTGTGTGCCTCTTCTTGCTGTTGCTGGCCCTT 1137

Qy     381 IleLeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIle 400
              |||:::|||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1138 GGACTCATTTACTGTCGCAAGAAGGAAGGGCTGGACTCCGATGTGGCCGACTCGTCCATC 1197

Qy     401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1198 CTCACCTCGGGCTTCCAGCCTGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCACCTG 1257

Qy     421 LeuThrIleGlnProAspLeuSer---ThrThrThrThrTyrGlnGlySerLeuCysPro 439
           ||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||   
Db    1258 CTCACCATCCAGCCAGACCTCAGCACCACCACTACCACCTACCAGGGCAGTCTATGTTCG 1317

Qy     440 ArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu 459
           |||||||||||||||||||||||||||||||||:::||||||||||||||||||||||||
Db    1318 AGGCAGGATGGACCCAGCCCCAAGTTCCAGCTCTCTAATGGTCACCTGCTCAGCCCACTG 1377

Qy     460 GlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheVal 479
           |||   |||||||||||||||||||||||||||||||||||||||||||:::||||||
Db    1378 GGGAGTGGCCGCCATACGTTGCACCACAGCTCACCCACCTCTGAGGCTGAGGACTTCGTC 1437

Qy     480 SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr 499
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||   
Db    1438 TCCCGCCTCTCCACCCAAAACTACTTTCGTTCCCTGCCCCGCGGCACCAGCAACATGGCC 1497

Qy     500 TyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeu 519
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1498 TACGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACGGGGATCAGCCTC 1557

Qy     520 LeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHis 539
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1558 CTCATACCCCCGGATGCCATCCCCCGAGGAAAGATCTACGAGATCTACCTCACACTGCAC 1617

Qy     540 LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal 559
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db    1618 AAGCCAGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCAGTCGTT 1677

Qy     560 SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys 579
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1678 AGCTGTGGGCCCCCAGGAGTCCTGCTCACCCGGCCAGTCATCCTTGCAATGGACCACTGT 1737

Qy     580 GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySer 599
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1738 GGAGAGCCCAGCCCTGACAGCTGGAGTCTGCGCCTCAAAAAGCAGTCCTGCGAGGGCAGT 1797

Qy     600 TrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGln 619
           ||||||   ||||||||||||||||||||||||:::||||||||||||||||||||||||
Db    1798 TGGGAG---GATGTGCTGCACCTTGGTGAGGAGTCACCTTCCCACCTCTACTACTGCCAG 1854

Qy     620 LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly 639
           |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db    1855 CTGGAGGCCGGGGCCTGCTATGTCTTCACGGAGCAGCTGGGCCGCTTTGCCCTGGTAGGA 1914

Qy     640 GluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAla 659
           |||||||||||||||||||||   |||||||||:::||||||||||||||||||||||||
Db    1915 GAGGCCCTCAGCGTGGCTGCCACCAAGCGCCTCAGGCTCCTTCTGTTTGCTCCCGTGGCC 1974

Qy     660 CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu 679
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1975 TGTACGTCCCTTGAGTACAACATCCGAGTGTACTGCCTACACGACACCCACGACGCTCTC 2034

Qy     680 LysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgVal 699
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2035 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTAGGTGGACAGCTGATCCAGGAGCCTCGCGTC 2094

Qy     700 LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2095 CTGCACTTCAAAGACAGTTACCACAACCTACGTCTCTCCATCCACGACGTGCCCAGCTCC 2154

Qy     720 LeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn 739
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2155 CTGTGGAAGAGCAAGCTACTTGTCAGCTACCAGGAGATCCCTTTTTACCACATCTGGAAC 2214

Qy     740 GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759
           |||||||||:::||||||||||||||||||||||||||||||::::::   |||||||||
Db    2215 GGCACCCAGCAGTATCTGCACTGCACCTTCACCCTGGAGCGCATCAACGCCAGCACCAGC 2274

Qy     760 AspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIle 779
           |||||||||||||||:::||||||||||||||||||||||||||||||||||||:::|||
Db    2275 GACCTGGCCTGCAAGGTGTGGGTGTGGCAGGTGGAGGGAGATGGGCAGAGCTTCAACATC 2334

Qy     780 AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAla 799
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||   
Db    2335 AACTTCAACATCACTAAGGACACAAGGTTTGCTGAATTGTTGGCTCTGGAGAGTGAAGGG 2394

Qy     800 GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2395 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAAAAG 2454

Qy     820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839
           ||||||:::||||||||||||||||||   ||||||||||||||||||||||||||||||
Db    2455 ATCATCGCCAGTCTGGACCCACCCTGCAGCCGGGGCGCCGACTGGAGAACTCTAGCCCAG 2514

Qy     840 LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2515 AAACTTCACCTGGACAGCCATCTTAGCTTCTTTGCCTCCAAGCCCAGCCCTACAGCCATG 2574

Qy     860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879
           |||||||||||||||||||||||||||||||||||||||||||||   ||||||||||||
Db    2575 ATCCTCAACCTATGGGAGGCACGGCACTTCCCCAACGGCAACCTCGGCCAGCTGGCAGCA 2634

Qy     880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2635 GCTGTGGCCGGACTGGGCCAACCAGATGCTGGCCTCTTCACGGTGTCGGAGGCCGAGTGT 2694
```

RESULT 3
US-08-808-982-2
; Sequence 2, Application US/08808982
; Patent No. 5939271
; GENERAL INFORMATION:

```
;    APPLICANT:  Tessier-Lavigne, Marc
;    APPLICANT:  Leonardo, E. David
;    APPLICANT:  Hink, Lindsay
;    APPLICANT:  Masu, Masayuki
;    APPLICANT:  Kazuko, Keino-Masu
;    TITLE OF INVENTION:  Netrin Receptors
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SCIENCE & TECHNOLOGY LAW GROUP
;      STREET:  268 BUSH STREET, SUITE 3200
;      CITY:  SAN FRANCISCO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/808,982
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  OSMAN, RICHARD A
;      REGISTRATION NUMBER:  36,627
;      REFERENCE/DOCKET NUMBER:  UC96-217
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 343-4341
;      TELEFAX:  (415) 343-4342
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1787 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-808-982-2

Alignment Scores:
Pred. No.:             8.57e-252    Length:         1787
Score:                 2694.00      Matches:        543
Percent Similarity:    97.14%       Conservative:   1
Best Local Similarity: 96.96%       Mismatches:     11
Query Match:           56.28%       Indels:         11
DB:                    2            Gaps:           2

US-09-970-944-2 (1-899) x US-08-808-982-2 (1-1787)

Qy        344 AsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspValAlaLeuTyr 363
              ||||||||||||||||||    ||||||:::|||||||||||||||||||||||||||||
Db          3 AACTGTACCAGTGACCTCTG-GTACACACTGCTTCTGGCCCTGAGGACGTGGCCCTCTAT 61

Qy        364 ValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuLeuValLeuIleLeuVal 383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 GTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCATCCTCGTT 121

Qy        384 TyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIleLeuThrSer 403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 TATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCA 181

Qy        404 GlyPheGlnProValSerIle-LysProSerLysAlaAspAsnProHisLeuLeuThrIl 423
              ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        182 GGCTTCCAGCCCGTCAGCATCTAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCAT 241

Qy        423 eGlnProAspLeuSer---ThrThrThrThrTyrGlnGlySerLeuCysProArgGlnAs 442
              ||||||||||||||||   |||||||||||||||||||||||||||||||||||||||||
Db        242 CCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGA 301

Qy        442 pGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeuGlyGlyGl 462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 TGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGG 361

Qy        462 yArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheValSerArgLe 482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 CCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCT 421

Qy        482 uSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThrTyrGlyTh 502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 CTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGAC 481

Qy        502 rPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeuLeuIlePr 522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 CTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGAATCAGCCTCCTCATCCC 541

Qy        522 oProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHisLysProGl 542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 CCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGA 601

Qy        542 uAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleValSerCysGl 562
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 AGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGG 661

Qy        562 yProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGlyGluPr 582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 ACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCC 721

Qy        582 oSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySerTrpGluGl 602
              |||||||||||||||||||||||    ||||||||||||||||||||||||   ||||||
Db        722 CAGCCCTGACAGCTGGAGCCTGGCCCTCAAAAAGCAGTCGTGCGAGGGAGC-TGGGAG-- 778

Qy        602 nAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGlnLeuGluAl 622
                |||    ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        779 -GATGT-CTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGC 836

Qy        622 aSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGluAlaLe 642
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        837 CAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCT 896

Qy        642 uSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAlaCysThrSe 662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        897 CAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTC 956

Qy        662 rLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuLysGluVa 682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        957 CCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGT 1016

Qy        682 lValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgValLeuHisPh 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1017 GGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTT 1076

Qy        702 eLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSerLeuTrpLy 722
               ||||||||||||||||||||    |||    |||||||||||||||||||||||||||||
Db       1077 -AAGGACAGTTACCACAACCTGCC-CTATC-ATCCACGATGTGCCCAGCTCCCTGTGGAA 1133

Qy        722 sSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsnGlyThrGl 742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1134 GAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCA 1193

Qy        742 nArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSerAspLeuAl 762
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1194 GCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGC 1253

Qy        762 aCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIleAsnPheAs 782
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1254 CTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAA 1313

Qy        782 nIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAlaGlyValPr 802
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1314 CATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCC 1373

Qy        802 oAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLysIleIleSe 822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1374 AGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTC 1433

Qy        822 rSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGlnLysLeuHi 842
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1434 CAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCA 1493

Qy         842 sLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMetIleLeuAs 862
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1494 CCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAA 1553

Qy         862 nLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAlaAlaValAl 882
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1554 CCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGC 1613

Qy         882 aGlyLeuGlyGlnProAspAlaGlyLeu-PheThrVal-SerGluAlaGluCys 899
               ||||  ||||||   |||   |||    ||||||||| ||||||||||||||
Db        1614 TGGGACTGGCCAGCAGGACGGTGGCTTCTTTCACAGTGTTCGGAGGCTGAGTGC 1667
```

RESULT 4
US-09-306-902A-2
; Sequence 2, Application US/09306902A
; Patent No. 6277585
;   GENERAL INFORMATION:
;        APPLICANT: Tessier-Lavigne, Marc
;                   Leonardo, E. David
;                   Hink, Lindsay
;                   Masu, Masayuki
;                   Kazuko, Keino-Masu
;        TITLE OF INVENTION: Netrin Receptors
;        NUMBER OF SEQUENCES: 9
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
;             STREET: 268 BUSH STREET, SUITE 3200
;             CITY: SAN FRANCISCO
;             STATE: CALIFORNIA
;             COUNTRY: USA
;             ZIP: 94104
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/306,902A
;             FILING DATE: 07-May-1999
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: OSMAN, RICHARD A
;             REGISTRATION NUMBER: 36,627
;             REFERENCE/DOCKET NUMBER: UC96-217
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 343-4341
;             TELEFAX: (415) 343-4342
;   INFORMATION FOR SEQ ID NO: 2:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1787 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-306-902A-2

```
Alignment Scores:
Pred. No.:               8.57e-252      Length:         1787
Score:                   2694.00        Matches:        543
Percent Similarity:      97.14%         Conservative:   1
Best Local Similarity:   96.96%         Mismatches:     11
Query Match:             56.28%         Indels:         11
DB:                      3              Gaps:           2

US-09-970-944-2 (1-899) x US-09-306-902A-2 (1-1787)

Qy         344 AsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspValAlaLeuTyr 363
               ||||||||||||||||||| |||||||:::||||||||||||||||||||||||||||||
Db           3 AACTGTACCAGTGACCTCTG-GTACACACTGCTTCTGGCCCTGAGGACGTGGCCCTCTAT 61

Qy         364 ValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuLeuValLeuIleLeuVal 383
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          62 GTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCATCCTCGTT 121

Qy         384 TyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIleLeuThrSer 403
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         122 TATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCA 181

Qy         404 GlyPheGlnProValSerIle-LysProSerLysAlaAspAsnProHisLeuLeuThrIl 423
               ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db         182 GGCTTCCAGCCCGTCAGCATCTAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCAT 241

Qy         423 eGlnProAspLeuSer---ThrThrThrThrTyrGlnGlySerLeuCysProArgGlnAs 442
               ||||||||||||||||   |||||||||||||||||||||||||||||||||||||||||
Db         242 CCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGA 301

Qy         442 pGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeuGlyGlyGl 462
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         302 TGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGG 361

Qy         462 yArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheValSerArgLe 482
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         362 CCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCT 421

Qy         482 uSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThrTyrGlyTh 502
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         422 CTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGAC 481

Qy         502 rPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeuLeuIlePr 522
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         482 CTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGAATCAGCCTCCTCATCCC 541

Qy         522 oProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHisLysProGl 542
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         542 CCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGA 601

Qy         542 uAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleValSerCysGl 562
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         602 AGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGG 661

Qy         562 yProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGlyGluPr 582
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         662 ACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCC 721

Qy         582 oSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySerTrpGluGl 602
               ||||||||||||||||||||      ||||||||||||||||||||||    |||||
Db         722 CAGCCCTGACAGCTGGAGCCTGGCCCTCAAAAAGCAGTCGTGCGAGGGAGC-TGGGAG-- 778

Qy         602 nAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGlnLeuGluAl 622
                |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         779 -GATGT-CTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGC 836

Qy         622 aSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGluAlaLe 642
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         837 CAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCT 896

Qy         642 uSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAlaCysThrSe 662
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         897 CAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTC 956

Qy         662 rLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuLysGluVa 682
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         957 CCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGT 1016

Qy         682 lValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgValLeuHisPh 702
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1017 GGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTT 1076

Qy         702 eLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSerLeuTrpLy 722
                ||||||||||||||||||||    |||    |||||||||||||||||||||||||||
Db        1077 -AAGGACAGTTACCACAACCTGCC-CTATC-ATCCACGATGTGCCCAGCTCCCTGTGGAA 1133
```

```
Qy        722 sSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsnGlyThrGl 742
Db       1134 GAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCA 1193

Qy        742 nArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSerAspLeuAl 762
Db       1194 GCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGC 1253

Qy        762 aCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIleAsnPheAs 782
Db       1254 CTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAA 1313

Qy        782 nIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAlaGlyValPr 802
Db       1314 CATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCC 1373

Qy        802 oAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLysIleIleSe 822
Db       1374 AGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTC 1433

Qy        822 rSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGlnLysLeuHi 842
Db       1434 CAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCA 1493

Qy        842 sLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMetIleLeuAs 862
Db       1494 CCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAA 1553

Qy        862 nLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAlaAlaValAl 882
Db       1554 CCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGC 1613

Qy        882 aGlyLeuGlyGlnProAspAlaGlyLeu-PheThrVal-SerGluAlaGluCys 899
Db       1614 TGGGACTGGCCAGCAGGACGGTGGCTTCTTTCACAGTGTTCGGAGGCTGAGTGC 1667

RESULT 5
US-08-808-982-3
; Sequence 3, Application US/08808982
; Patent No. 5939271
;  GENERAL INFORMATION:
;    APPLICANT:  Tessier-Lavigne, Marc
;    APPLICANT:  Leonardo, E. David
;    APPLICANT:  Hink, Lindsay
;    APPLICANT:  Masu, Masayuki
;    APPLICANT:  Kazuko, Keino-Masu
;    TITLE OF INVENTION:  Netrin Receptors
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SCIENCE & TECHNOLOGY LAW GROUP
;      STREET:  268 BUSH STREET, SUITE 3200
;      CITY:  SAN FRANCISCO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/808,982
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  OSMAN, RICHARD A
;      REGISTRATION NUMBER:  36,627
;      REFERENCE/DOCKET NUMBER:  UC96-217
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 343-4341
;      TELEFAX:  (415) 343-4342
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2831 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-808-982-3

Alignment Scores:
Pred. No.:              5.72e-229    Length:         2831
Score:                  2463.00      Matches:        495
Percent Similarity:     67.41%       Conservative:   144
Best Local Similarity:  52.22%       Mismatches:     227
Query Match:            51.45%       Indels:         83
DB:                     2            Gaps:           19

US-09-970-944-2 (1-899) x US-08-808-982-3 (1-2831)

Qy          9 ProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeuArgGlySerGlyAlaGlnGlnSer 28
Db         61 CCGAGCTTAGCAGGCATT--------------------GACTCTGGTGCCCAG------ 93

Qy         29 AlaThrValAlaAsnProValProGlyAlaAsnProAspLeuLeuProHisPheLeuVal 48
Db         94 ---GGACTCCCAGACTCCTTCCCATCAGCACCCGCGGAGCAGCTGCCTCACTTCCTGCTG 150

Qy         49 GluProGluAspValTyrIleValLysAsnLysProValLeuLeuValCysLysAlaVal 68
Db        151 GAACCAGAGGATGCCTACATCGTAAAGAACAAGCCAGTGGAATTGCACTGCCGAGCCTTC 210

Qy         69 ProAlaThrGlnIlePhePheLysCysAsnGlyGluTrpValArgGlnValAspHisVal 88
Db        211 CCTGCCACACAGATCTACTTCAAGTGTAATGGCGAGTGGGTTAGCCAGAAAGGCCACGTC 270

Qy         89 IleGluArgSerThrAspGlySerSerGlyGluProThrMetGluValArgIleAsnVal 108
Db        271 ACGCAGGAGAGCCTGGATGAGGCCACAGGCTTGCGAATACGAGAGGTGCAGATAGAGGTG 330

Qy        109 SerArgGlnGlnValGluLysValPheGlyLeuGluGluTyrTrpCysGlnCysValAla 128
Db        331 TCGCGGCAGCAGGTGGAGGAACTTTTTGGGCTCGAGGACTACTGGTGTCAGTGCGTGGCC 390

Qy        129 TrpSerSerSerGlyThrThrLysSerGlnLysAlaTyrIleArgIleAlaArgLeuArg 148
Db        391 TGGAGCTCTTCGGGAACCACCAAGAGTCGCCGAGCCTACATCCGCATTGCCTACTTGCGC 450

Qy        149 LysAsnPheGluGlnGluProLeuAlaLysGluValSerLeuGluGlnGlyIleValLeu 168
Db        451 AAGAACTTTGACCAGGAGCCTCTGGCGAAGGAGGTACCCTTGGATCATGAGGTCCTTCTG 510

Qy        169 ProCysArgProProGluGlyIleProProAlaGluValGluTrpLeuArgAsnGluAsp 188
Db        511 CAGTGCCGCCCACCAGAGGGAGTGCCTGTGGCTGAGGTGGAATGGCTCAAGAATGAAGAT 570

Qy        189 LeuValAspProSerLeuAspProAsnValTyrIleThrArgGluHisSerLeuValVal 208
Db        571 GTCATCGATCCCGCTCAGGACACTAACTTCCTGCTCACCATTGACCACAACCTCATCATC 630

Qy        209 ArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThrCysValAlaLysAsnIleValAla 228
Db        631 CGCCAGGCGCGCCTCTCAGACACAGCCAACTACACCTGTGTGGCAAAGAATATTGTGGCC 690

Qy        229 ArgArgArgSerAlaSerAlaAlaValIleValTyrValAsnGlyGlyTrpSerThrTrp 248
Db        691 AAGCGCCGGAGCACGACGGCCACAGTCATCGTCTATGTGAACGGAGGTTGGTCCAGCTGG 750

Qy        249 ThrGluTrpSerValCysSerAlaSerCysGlyArgGlyTrpGlnLysArgSerArgSer 268
Db        751 GCAGAATGGTCACCCTGCTCTAACCGCTGCGGCCGAGGTTGGCAGAAACGTACTAGGACC 810

Qy        269 CysThrAsnProAlaProLeuAsnGlyGlyAlaPheCysGluGlyGlnAsnValHisAsp 288
Db        811 TGCACCAACCCAGCCCCACTCAATGGAGGTGCCTTCTGCGAGGGACAG---GCTTGCCAG 867
```

```
Qy   289 ArgThrValSerSerLeuLeuValSerValAspGlySerTrpSerProTrpSerLysTrp 308
Db   868 AAGACGGCTTGCACCACCGTGTGCCCAGTGGATGGAGCGTGGACTGAGTGGAGCAAGTGG 927

Qy   309 SerAlaCysGlyLeuAspCysThrHisTrpArgSerArgGluCysSerAspProAlaPro 328
Db   928 TCCGCCTGCAGCACAGAGTGTGCGCACTGGCGCAGCCGCGAGTGCATGGCACCGCCGCCC 987

Qy   329 ArgAsnGlyGlyGluGluCysGlnGlyThrAspLeuAspThrArgAsnCysThrSerAsp 348
Db   988 CAGAACGGAGGCCGTGACTGCAGCGGGACGCTACTTGACTCCAAGAACTGCACCGATGGG 1047

Qy   349 LeuCysVal---------HisSerAlaSerGlyProGlu--------------------- 358
Db  1048 CTGTGCGTGCTGAATCAGAGAACTCTAAACGACCCTAAAAGCCGCCCCCTGGAGCCGTCG 1107

Qy   359 ---AspValAlaLeuTyrValGlyLeu---IleAlaValAlaValCysLeuValLeuLeu 376
Db  1108 GGAGACGTGGCGCTGTATGCGGGCCTCGTGGTGGCCGTCTTTGTGGTTCTGGCAGTTCTC 1167

Qy   377 LeuLeuValLeuIleLeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAla 396
Db  1168 ATGGCTGTAGGAGTGATCGTGTACCGGAGAAACTGCCGGGACTTCGACACGGACATCACT 1227

Qy   397 AspSerSer---IleLeuThrSerGlyPheGlnProValSerIleLysProSerLysAla 415
Db  1228 GACTCCTCTGCTGCCCTCACTGGTGGTTTCCACCCCGTCAACTTCAAGACTGCAAGGCCC 1287

Qy   416 AspAsnProHisLeuLeu------ThrIleGlnProAspLeuSerThrThrThr---Thr 432
Db  1288 AGCAACCCACAGCTCCTGCACCCATCCGCCCCTCCGGACCTAACGGCCAGTGCTGGCATC 1347

Qy   433 TyrGlnGlySerLeuCysProArgGlnAspGlyProSerProLysPheGlnLeuThrAsn 452
Db  1348 TACCGCGGACCTGTGTATGCCCTGCAGGACTCT---GCCGACAAGATCCCTATGACTAAT 1404

Qy   453 GlyHisLeuLeuSerProLeu--------------------------------------- 459
Db  1405 TCACCCCTTCTGGATCCCTTGCCCAGCCTCAAGATCAAGGTCTATGACTCCAGCACCATC 1464

Qy   460 ------GlyGlyGly--------------------------------------------- 462
Db  1465 GGCTCTGGGGCTGGCCTGGCTGATGGAGCCGACCTGCTGGGTGTCTTACCACCCGGTACA 1524

Qy   463 ------------------------ArgHisThrLeuHisHisSerSerProThrSerGlu 474
Db  1525 TACCCAGGCGATTTCTCCCGGGACACCCACTTCCTGCACCTGCGCAGC------------ 1572

Qy   475 AlaGluGluPheValSerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGly 494
Db  1573 ---------------GCCAGCCTTGGTTCCCAG---CACCTCCTGGGCCTCCCTCGAGAC 1614

Qy   495 ThrSerAsnMetThrTyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsn 514
Db  1615 CCCAGCAGCAGTGTCAGTGGCACCTTTGGTTGCCTGGGTGGAGGCTGACCATTCCCGGC 1674

Qy   515 ThrGlyIleSerLeuLeuIleProProAspAlaIleProArgGlyLysIleTyrGluIle 534
Db  1675 ACAGGGGTCAGCCTGTTGGTACCAAATGGAGCCATTCCCCAGGGCAAGTTCTATGACTTG 1734

Qy   535 TyrLeuThrLeuHisLysProGluAspValArgLeuProLeuAla---GlyCysGlnThr 553
Db  1735 TATCTACGTATCAACAAGACTGAAAGCACC---CTCCCACTTTCGGAAGGTTCCCAGACA 1791

Qy   554 LeuLeuSerProIleValSerCysGlyProProGlyValLeuLeuThrArgProValIle 573
Db  1792 GTATTGAGCCCCTCGGTGACCTGCGGGCCCACGGGCCTCCTCCTGTGCCGCCCTGTTGTC 1851

Qy   574 LeuAlaMetAspHisCysGlyGluProSerProAspSerTrpSerLeuArgLeuLysLys 593
Db  1852 CTCACTGTGCCCCACTGTGCTGAAGTCATTGCCGGAGACTGGATCTTCCAGCTCAAGACC 1911

Qy   594 GlnSerCysGluGlySerTrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSer 613
Db  1912 CAGGCCCATCAGGGCCACTGGGAG---GAGGTGGTGACTTTGGATGAGGAGACTCTGAAC 1968

Qy   614 HisLeuTyrTyrCysGlnLeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGly 633
Db  1969 ACCCCCTGCTACTGCCAGCTAGAGGCTAAATCCTGCCACATCCTGTTGGACCAGCTGGGT 2028

Qy   634 ArgPheAlaLeuValGlyGluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeu 653
Db  2029 ACCTACGTGTTCACGGGCGAGTCCTACTCCCGCTCCGCAGTCAAGCGGCTCCAGCTAGCC 2088

Qy   654 LeuPheAlaProValAlaCysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHis 673
Db  2089 ATCTTCGCCCCAGCCCTCTGCACCTCCCTGGAGTATAGTCTCAGGGTCTACTGTCTGGAG 2148

Qy   674 AspThrHisAspAlaLeuLysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeu 693
Db  2149 GACACTCCTGCAGCACTGAAGGAGGTCCTAGAGCTGGAGAGGACTCTGGGTGGCTACTTG 2208

Qy   694 IleGlnGluProArgValLeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIle 713
Db  2209 GTGGAGGAGCCCAAGACTTTGCTCTTTAAGGACAGTTACCACAACCTACGCTC-TCCCTC 2267

Qy   714 HisAspValProSerSerLeuTrpLysSerLysLeuLeuValSerTyrGlnGluIlePro 733
Db  2268 CATGACATCCCCCATGCCCACTGGAGGAGCAAACTACTGGCCAAGTACCAGGAGATTCCC 2327

Qy   734 PheTyrHisIleTrpAsnGlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArg 753
Db  2328 TTCTACCATGTGTGGAACGGCAGCCAGAAAGCCCTGCACTGCACTTTCACCCTGGAGAGA 2387

Qy   754 ValSerProSerThrSerAspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAsp 773
Db  2388 CATAGCCTAGCCTCCACTGAGTTCACCTGTAAGGTCTGCGTGCGGCAGGTAGAAGGGGAA 2447

Qy   774 GlyGlnSerPheSerIleAsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeu 793
Db  2448 GGCCAGATTTTCCAGCTGCACACCACGCTGGCT---GAGACGCCTGCTGGCTCCCTGGAT 2504

Qy   794 AlaLeuGluSerGluAlaGlyValProAlaLeu------ValGlyProSerAlaPheLys 811
Db  2505 GCACTCTGCTCTGCCCCTGGCAATGCTGCCACCACACAGCTGGGACCCTATGCCTTCAAG 2564

Qy   812 IleProPheLeuIleArgGlnLysIleIleSerSerLeuAspProProCysArgArgGly 831
Db  2565 ATACCACTGTCCATCCGCCAGAAGATCTGCAACAGCCTGGACGCCCCCAACTCACGGGGC 2624

Qy   832 AlaAspTrpArgThrLeuAlaGlnLysLeuHisLeuAspSerHisLeuSerPhePheAla 851
Db  2625 AATGACTGGCGGCTGTTGGCACAGAAGCTCTCCATGGACCGGTACCTGAACTACTTCGCC 2684

Qy   852 SerLysProSerProThrAlaMetIleLeuAsnLeuTrpGluAlaArgHisPheProAsn 871
Db  2685 ACCAAAGCTAGTCCCACAGGCGTGATCTTAGACCTCTGGGAAGCTCGGCAGCAGGATGAT 2744

Qy   872 GlyAsnLeuSerGlnLeuAlaAlaAlaValAlaGlyLeuGlyGlnProAspAlaGlyLeu 891
Db  2745 GGGGACCTCAACAGCCTGGCCAGTGCCTTGGAGGAGATGGGCAAGAGTGAGATGCTGGTA 2804

Qy   892 PheThrValSerGluAlaGluCys 899
Db  2805 GCCATGACCACTGATGGCGATTGC 2828
```

RESULT 6
US-09-306-902A-3
; Sequence 3, Application US/09306902A
; Patent No. 6277585
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; Leonardo, E. David
; Hink, Lindsay
; Masu, Masayuki
; Kazuko, Keino-Masu

```
;       TITLE OF INVENTION: Netrin Receptors
;       NUMBER OF SEQUENCES: 9
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
;            STREET: 268 BUSH STREET, SUITE 3200
;            CITY: SAN FRANCISCO
;            STATE: CALIFORNIA
;            COUNTRY: USA
;            ZIP: 94104
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/306,902A
;            FILING DATE: 07-May-1999
;            CLASSIFICATION: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: OSMAN, RICHARD A
;            REGISTRATION NUMBER: 36,627
;            REFERENCE/DOCKET NUMBER: UC96-217
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (415) 343-4341
;            TELEFAX: (415) 343-4342
;  INFORMATION FOR SEQ ID NO: 3:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 2831 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: double
;            TOPOLOGY: linear
;       MOLECULE TYPE: cDNA
;       SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-306-902A-3

Alignment Scores:
Pred. No.:              5.72e-229     Length:         2831
Score:                  2463.00       Matches:        495
Percent Similarity:     67.41%        Conservative:   144
Best Local Similarity:  52.22%        Mismatches:     227
Query Match:            51.45%        Indels:         83
DB:                     3             Gaps:           19

US-09-970-944-2 (1-899) x US-09-306-902A-3 (1-2831)

Qy          9 ProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeuArgGlySerGlyAlaGlnGlnSer 28
              |||:::|||   |||||                     |||||||||||
Db         61 CCGAGCTTAGCAGGCATT--------------------GACTCTGGTGCCCAG------ 93

Qy         29 AlaThrValAlaAsnProValProGlyAlaAsnProAspLeuLeuProHisPheLeuVal 48
                 :::   :::        |||   |||     :::   |||||||||||||||:::
Db         94 ---GGACTCCCAGACTCCTTCCCATCAGCACCCGCGGAGCAGCTGCCTCACTTCCTGCTG 150

Qy         49 GluProGluAspValTyrIleValLysAsnLysProValLeuLeuValCysLysAlaVal 68
              |||||||||||   ||||||||||||||||||||||||    |||    |||:::|||
Db        151 GAACCAGAGGATGCCTACATCGTAAAGAACAAGCCAGTGGAATTGCACTGCCGAGCCTTC 210

Qy         69 ProAlaThrGlnIlePhePheLysCysAsnGlyGluTrpValArgGlnValAspHisVal 88
              ||||||||||||||:::||||||||||||||||||||||    |||      ||||||
Db        211 CCTGCCACACAGATCTACTTCAAGTGTAATGGCGAGTGGGTTAGCCAGAAAGGCCACGTC 270

Qy         89 IleGluArgSerThrAspGlySerSerGlyGluProThrMetGluValArgIleAsnVal 108
                 :::   |||   |||   ::::::|||            ||||||:::|||   |||
Db        271 ACGCAGGAGAGCCTGGATGAGGCCACAGGCTTGCGAATACGAGAGGTGCAGATAGAGGTG 330

Qy        109 SerArgGlnGlnValGluLysValPheGlyLeuGluGluTyrTrpCysGlnCysValAla 128
              ||||||||||||||||||::::::||||||||||||:::|||||||||||||||||||||
Db        331 TCGCGGCAGCAGGTGGAGGAACTTTTGGGCTCGAGGACTACTGGTGTCAGTGCGTGGCC 390

Qy        129 TrpSerSerSerGlyThrThrLysSerGlnLysAlaTyrIleArgIleAlaArgLeuArg 148
              ||||||||||||||||||||||||::::::||||||||||||||||||    ||||||
Db        391 TGGAGCTCTTCGGGAACCACCAAGAGTCGCCGAGCCTACATCCGCATTGCCTACTTGCGC 450

Qy        149 LysAsnPheGluGlnGluProLeuAlaLysGluValSerLeuGluGlnGlyIleValLeu 168
              ||||||||:::|||||||||||||||||||||||||    |||:::      ::::::|||
Db        451 AAGAACTTTGACCAGGAGCCTCTGGCGAAGGAGGTACCCTTGGATCATGAGGTCCTTCTG 510

Qy        169 ProCysArgProProGluGlyIleProProAlaGluValGluTrpLeuArgAsnGluAsp 188
                 ||||||||||||||||:::|||   ||||||||||||||||||:::||||||||
Db        511 CAGTGCCGCCCACCAGAGGGAGTGCCTGTGGCTGAGGTGGAATGGCTCAAGAATGAAGAT 570

Qy        189 LeuValAspProSerLeuAspProAsnValTyrIleThrArgGluHisSerLeuValVal 208
              ::::::||||||:::      |||   |||      :::|||   :::|||:::|||::::::
Db        571 GTCATCGATCCCGCTCAGGACACTAACTTCCTGCTCACCATTGACCACAACCTCATCATC 630

Qy        209 ArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThrCysValAlaLysAsnIleValAla 228
              |||||||||||||||:::||||||||||||||||||||||||||||||||||||||||||
Db        631 CGCCAGGCGCGCCTCTCAGACACAGCCAACTACACCTGTGTGGCAAAGAATATTGTGGCC 690

Qy        229 ArgArgArgSerAlaSerAlaAlaValIleValTyrValAsnGlyGlyTrpSerThrTrp 248
              :::|||||||||      :::|||   ||||||||||||||||||||||||||||:::|||
Db        691 AAGCGCCGGAGCACGACGGCCACAGTCATCGTCTATGTGAACGGAGGTTGGTCCAGCTGG 750

Qy        249 ThrGluTrpSerValCysSerAlaSerCysGlyArgGlyTrpGlnLysArgSerArgSer 268
                 ||||||||   ||||||      |||||||||||||||||||||||:::|||:::
Db        751 GCAGAATGGTCACCCTGCTCTAACCGCTGCGGCCGAGGTTGGCAGAAACGTACTAGGACC 810

Qy        269 CysThrAsnProAlaProLeuAsnGlyGlyAlaPheCysGluGlyGlnAsnValHisAsp 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        811 TGCACCAACCCAGCCCCACTCAATGGAGGTGCCTTCTGCGAGGGACAG---GCTTGCCAG 867

Qy        289 ArgThrValSerSerLeuLeuValSerValAspGlySerTrpSerProTrpSerLysTrp 308
              :::|||      :::   :::         |||||||||:::|||:::   ||||||||||
Db        868 AAGACGGCTTGCACCACCGTGTGCCCAGTGGATGGAGCGTGGACTGAGTGGAGCAAGTGG 927

Qy        309 SerAlaCysGlyLeuAspCysThrHisTrpArgSerArgGluCysSerAspProAlaPro 328
              ||||||||      :::|||   ||||||||||||||||||      |||   |||
Db        928 TCCGCCTGCAGCACAGAGTGTGCGCACTGGCGCAGCCGCGAGTGCATGGCACCGCCGCCC 987

Qy        329 ArgAsnGlyGlyGluGluCysGlnGlyThrAspLeuAspThrArgAsnCysThrSerAsp 348
              :::||||||||      :::|||   ||||||   ||||||::::::|||||||||
Db        988 CAGAACGGAGGCCGTGACTGCAGCGGGACGCTACTTGACTCCAAGAACTGCACCGATGGG 1047

Qy        349 LeuCysVal---------HisSerAlaSerGlyProGlu-------------------- 358
              |||||||||           :::   :::   |||:::
Db       1048 CTGTGCGTGCTGAATCAGAGAACTCTAAACGACCCTAAAAGCCGCCCCCTGGAGCCGTCG 1107

Qy        359 ---AspValAlaLeuTyrValGlyLeu---IleAlaValAlaValCysLeuValLeuLeu 376
                 ||||||||||||||   ||||||   :::||||||   |||   |||   :::|||
Db       1108 GGAGACGTGGCGCTGTATGCGGGCCTCGTGGTGGCCGTCTTTGTGGTTCTGGCAGTTCTC 1167

Qy        377 LeuLeuValLeuIleLeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAla 396
              :::   |||   ::::::||||||   |||               |||:::|||:::
Db       1168 ATGGCTGTAGGAGTGATCGTGTACCGGAGAAACTGCCGGGACTTCGACACGGACATCACT 1227

Qy        397 AspSerSer---IleLeuThrSerGlyPheGlnProValSerIleLysProSerLysAla 415
              ||||||||         ||||||   ||||||   ||||||:::   |||   ::::::
Db       1228 GACTCCTCTGCTGCCCTCACTGGTGGTTTCCACCCCGTCAACTTCAAGACTGCAAGGCCC 1287

Qy        416 AspAsnProHisLeuLeu------ThrIleGlnProAspLeuSerThrThrThr---Thr 432
               ||||||   ||||||         :::      |||||||||:::   :::
Db       1288 AGCAACCCACAGCTCCTGCACCCATCCGCCCCTCCGGACCTAACGGCCAGTGCTGGCATC 1347

Qy        433 TyrGlnGlySerLeuCysProArgGlnAspGlyProSerProLysPheGlnLeuThrAsn 452
              |||:::|||      :::         ||||||      :::   |||      :::||||||
Db       1348 TACCGCGGACCTGTGTATGCCCTGCAGGACTCT---GCCGACAAGATCCCTATGACTAAT 1404

Qy        453 GlyHisLeuLeuSerProLeu-------------------------------------- 459
                    ||||||   ||||||
Db       1405 TCACCCCTTCTGGATCCCTTGCCCAGCCTCAAGATCAAGGTCTATGACTCCAGCACCATC 1464

Qy        460 ------GlyGlyGly-------------------------------------------- 462
                    |||   |||
Db       1465 GGCTCTGGGGCTGGCCTGGCTGATGGAGCCGACCTGCTGGGTGTCTTACCACCCGGTACA 1524

Qy        463 -----------------------ArgHisThrLeuHisHisSerSerProThrSerGlu 474
```

```
Db       1525 TACCCAGGCGATTTCTCCCGGGACACCCACTTCCTGCACCTGCGCAGC------------ 1572

Qy        475 AlaGluGluPheValSerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGly 494
Db       1573 ---------------GCCAGCCTTGGTTCCCAG---CACCTCCTGGGCCTCCCTCGAGAC 1614

Qy        495 ThrSerAsnMetThrTyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsn 514
Db       1615 CCCAGCAGCAGTGTCAGTGGCACCTTTGGTTGCCTGGGTGGGAGGCTGACCATTCCCGGC 1674

Qy        515 ThrGlyIleSerLeuLeuIleProProAspAlaIleProArgGlyLysIleTyrGluIle 534
Db       1675 ACAGGGGTCAGCCTGTTGGTACCAAATGGAGCCATTCCCCAGGGCAAGTTCTATGACTTG 1734

Qy        535 TyrLeuThrLeuHisLysProGluAspValArgLeuProLeuAla---GlyCysGlnThr 553
Db       1735 TATCTACGTATCAACAAGACTGAAAGCACC---CTCCCACTTTCGGAAGGTTCCCAGACA 1791

Qy        554 LeuLeuSerProIleValSerCysGlyProProGlyValLeuLeuThrArgProValIle 573
Db       1792 GTATTGAGCCCCTCGGTGACCTGCGGGCCCACGGGCCTCCTCCTGTGCCGCCCTGTTGTC 1851

Qy        574 LeuAlaMetAspHisCysGlyGluProSerProAspSerTrpSerLeuArgLeuLysLys 593
Db       1852 CTCACTGTGCCCCACTGTGCTGAAGTCATTGCCGGAGACTGGATCTTCCAGCTCAAGACC 1911

Qy        594 GlnSerCysGluGlySerTrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSer 613
Db       1912 CAGGCCCATCAGGGCCACTGGGAG---GAGGTGGTGACTTTGGATGAGGAGACTCTGAAC 1968

Qy        614 HisLeuTyrTyrCysGlnLeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGly 633
Db       1969 ACCCCCTGCTACTGCCAGCTAGAGGCTAAATCCTGCCACATCCTGTTGGACCAGCTGGGT 2028

Qy        634 ArgPheAlaLeuValGlyGluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeu 653
Db       2029 ACCTACGTGTTCACGGGCGAGTCCTACTCCCGCTCCGCAGTCAAGCGGCTCCAGCTAGCC 2088

Qy        654 LeuPheAlaProValAlaCysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHis 673
Db       2089 ATCTTCGCCCCAGCCCTCTGCACCTCCCTGGAGTATAGTCTCAGGGTCTACTGTCTGGAG 2148

Qy        674 AspThrHisAspAlaLeuLysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeu 693
Db       2149 GACACTCCTGCAGCACTGAAGGAGGTCCTAGAGCTGGAGAGGACTCTGGGTGGCTACTTG 2208

Qy        694 IleGlnGluProArgValLeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIle 713
Db       2209 GTGGAGGAGCCCAAGACTTTGCTCTTTAAGGACAGTTACCACAACCTACGCTC-TCCCTC 2267

Qy        714 HisAspValProSerSerLeuTrpLysSerLysLeuLeuValSerTyrGlnGluIlePro 733
Db       2268 CATGACATCCCCCATGCCCACTGGAGGAGCAAACTACTGGCCAAGTACCAGGAGATTCCC 2327

Qy        734 PheTyrHisIleTrpAsnGlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArg 753
Db       2328 TTCTACCATGTGTGGAACGGCAGCCAGAAAGCCCTGCACTGCACTTTCACCCTGGAGAGA 2387

Qy        754 ValSerProSerThrSerAspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAsp 773
Db       2388 CATAGCCTAGCCTCCACTGAGTTCACCTGTAAGGTCTGCGTGCGGCAGGTAGAAGGGGAA 2447

Qy        774 GlyGlnSerPheSerIleAsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeu 793
Db       2448 GGCCAGATTTTCCAGCTGCACACCACGCTGGCT---GAGACGCCTGCTGGCTCCCTGGAT 2504

Qy        794 AlaLeuGluSerGluAlaGlyValProAlaLeu------ValGlyProSerAlaPheLys 811
Db       2505 GCACTCTGCTCTGCCCCTGGCAATGCTGCCACCACACAGCTGGGACCCTATGCCTTCAAG 2564

Qy        812 IleProPheLeuIleArgGlnLysIleIleSerSerLeuAspProProCysArgArgGly 831
Db       2565 ATACCACTGTCCATCCGCCAGAAGATCTGCAACAGCCTGGACGCCCCCAACTCACGGGGC 2624

Qy        832 AlaAspTrpArgThrLeuAlaGlnLysLeuHisLeuAspSerHisLeuSerPhePheAla 851
Db       2625 AATGACTGGCGGCTGTTGGCACAGAAGCTCTCCATGGACCGGTACCTGAACTACTTCGCC 2684

Qy        852 SerLysProSerProThrAlaMetIleLeuAsnLeuTrpGluAlaArgHisPheProAsn 871
Db       2685 ACCAAAGCTAGTCCCACAGGCGTGATCTTAGACCTCTGGGAAGCTCGGCAGCAGGATGAT 2744

Qy        872 GlyAsnLeuSerGlnLeuAlaAlaAlaValAlaGlyLeuGlyGlnProAspAlaGlyLeu 891
Db       2745 GGGGACCTCAACAGCCTGGCCAGTGCCTTGGAGGAGATGGGCAAGAGTGAGATGCTGGTA 2804

Qy        892 PheThrValSerGluAlaGluCys 899
Db       2805 GCCATGACCACTGATGGCGATTGC 2828
```

```
RESULT 7
US-09-833-381-1806
; Sequence 1806, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
;  APPLICANT: Robison, Keith E.
;  TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
;  FILE REFERENCE: 5800-119
;  CURRENT APPLICATION NUMBER: US/09/833,381
;  CURRENT FILING DATE:  2001-04-11
;  PRIOR APPLICATION NUMBER: 09/516,448
;  PRIOR FILING DATE: 2000-02-29
;  NUMBER OF SEQ ID NOS: 2050
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1806
;   LENGTH: 1282
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(1282)
;   OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1806

Alignment Scores:
Pred. No.:               2.49e-188      Length:         1282
Score:                   2039.00        Matches:        410
Percent Similarity:      95.58%         Conservative:   1
Best Local Similarity:   95.35%         Mismatches:     11
Query Match:             42.59%         Indels:         9
DB:                      4              Gaps:           2

US-09-970-944-2 (1-899) x US-09-833-381-1806 (1-1282)

Qy        476 GluGluPheValSerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThr 495
Db          3 GAGGAGTTCGTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACC 62

Qy        496 SerAsnMetThrTyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThr 515
Db         63 AGCAACATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACA 122

Qy        516 GlyIleSerLeuLeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyr 535
Db        123 GGAATCAGCCTCCTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTAC 182

Qy        536 LeuThrLeuHisLysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeu 555
Db        183 CTCACGCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTG 242

Qy        556 SerProIleValSerCysGlyProPro-GlyValLeuLeuThrArgProValIleLeu-A 575
Db        243 AGTCCCATCGTTAGCTGTGGACCCCCTGGGCGTCCTGCTCACCCGGCCAGTCATCCTGGG 302

Qy        575 laMetAspHisCysGlyGluProSerProAspSerTrp-SerLeuArgLeuLysLysGln 594
```

```
Db        303 GTATGGACCACTGTGGGGAGCCCAGCCCTGACAGCTGGGAGCCTGCGCCTCAAAAAGCAG 362

Qy        595 SerCysGluGlySerTrpGluGlnAspVal-LeuHisLeuGlyGluGluAlaProSerHi 614
Db        363 TCGTGCGAGGGCAGCTGGGAG---GATGTGCTTGCACCTGGGCGAGGAGGCGCCCTCCCA 419

Qy        614 sLeuTyrTyrCysGlnLeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyAr 634
Db        420 CCTCTACTACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGAGCCG 479

Qy        634 gPheAlaLeuValGlyGluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLe 654
Db        480 CTATGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCT 539

Qy        654 uPheAlaProValAlaCysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAs 674
Db        540 GTTTGCGCCGGTGGCCTGCACCTCCCTCGAGTACAACATACTGGTCTACTGCCTGCATGA 599

Qy        674 pThrHisAspAlaLeuLysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIl 694
Db        600 CACTCACGATGCACTCAACGTAGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGAT 659

Qy        694 eGlnGluProArgValLeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHi 714
Db        660 CCAGGAGCCACGGGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCA 719

Qy        714 sAspValProSerSerLeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPh 734
Db        720 CGATGTGCCCAGCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTT 779

Qy        734 eTyrHisIleTrpAsnGlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgVa 754
Db        780 TTATCACATCTGGAATGGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGT 839

Qy        754 lSerProSerThrSerAspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGl 774
Db        840 CAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGG 899

Qy        774 yGlnSerPheSerIleAsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAl 794
Db        900 GCAGAGCTTCAGCATCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGC 959

Qy        794 aLeuGluSerGluAlaGlyValProAlaLeuValGlyProSerAlaPheLysIleProPh 814
Db        960 TCTGGAGAGTGAAGCGNGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTT 1019

Qy        814 eLeuIleArgGlnLysIleIle-SerSerLeuAspProProCysArgArgGlyAlaAspT 834
Db       1020 CCTCATTCGGCAGAAGATAATTTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACT 1079

Qy        834 rpArg-ThrLeuAlaGlnLysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLys 853
Db       1080 GGCGGAACTCTGGCCCAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAG 1139

Qy        854 ProSerProThrAlaMetIleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsn 873
Db       1140 CCCAGCCCCACAGCCATGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAAC 1199

Qy        874 LeuSerGlnLeuAlaAlaAlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThr 893
Db       1200 CTCAGCCAGCTGGT-GCAGCAGTGGCTGGACTGGGCAGC---AGACGCTGGCTCTTCACA 1255

Qy        894 ValSerGluAlaGluCys 899
Db       1256 GTGTCGGAGGCTGAGTGC 1273
```

RESULT 8
US-09-833-381-1807
; Sequence 1807, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1807
; LENGTH: 1605
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1605)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1807

Alignment Scores:

| | | | |
|---|---|---|---|
| Pred. No.: | 4.72e-99 | Length: | 1605 |
| Score: | 1123.00 | Matches: | 254 |
| Percent Similarity: | 62.38% | Conservative: | 81 |
| Best Local Similarity: | 47.30% | Mismatches: | 155 |
| Query Match: | 23.46% | Indels: | 47 |
| DB: | 4 | Gaps: | 12 |

US-09-970-944-2 (1-899) x US-09-833-381-1807 (1-1605)

```
Qy        385 CysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSer---IleLeuThrSer 403
Db         19 TGCCGT---------GACTTCGACACAGACATCACTGACTCATCTGCTGCCCTGACTGGT 69

Qy        404 GlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeuLeu------ 421
Db         70 GGTTTCCACCCCGTCAACTTTAAGACGGCAAGGCCCAGCAACCCGCAGCTCCTACACCCC 129

Qy        422 ThrIleGlnProAspLeuSerThrThrThr---ThrTyrGlnGlySerLeuCysProArg 440
Db        130 TCTGTGCCTCCTGACCTGACAGCCAGCGCCGGCATCTACCGCGGACCCGTGTATGCCCTG 189

Qy        441 GlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeuGly 460
Db        190 CAGGACTCC---ACCGACAAAATCCCCATGACCAACTCTCCTCTGCTGGACCCCTTACCC 246

Qy        461 GlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAla--------------- 475
Db        247 AGCCTTAAGGTCAAGGTCCACACCTCCAGCACCACGGGCTCTGGGCCAGGCCTGGCAGAT 306

Qy        475 ------------------------------------------------------------ 475
Db        307 GGGGCTGACCTGCTGGGGGTCTTGCCGCCTGGCACATACCCTAGCGATTTCACCCGGGAC 366

Qy        476 GluGluPheValSerArgLeuSerThrGlnAsnTyrPhe-Arg---------SerLeuPr 492
Db        367 ACCCACTTCCTGCACCTGCGCAGCCGCCAGCCTCGGTTCCCAGCAGCTCTTGGGCCTGCC 426

Qy        492 oArgGlyThrSerAsnMetThrTyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIl 512
Db        427 CCGAGACCCAGGGAGCAGCGTCAGCGGCACCTTTGGCTGCCTGGGTGGAGGCTCANCAT 486

Qy        512 eProAsnThrGlyIleSerLeuLeuIleProProAspAlaIleProArgGlyLysIleTy 532
Db        487 CCCCGGCACAGGGGTCAGCTTGCTGGTGCCCAATGGAGCCATTCCCCAGGGCAAGTTCTA 546

Qy        532 rGlu-IleTyr-LeuThrLeuHisLysProGluAspValArgLeuProLeuAla---Gly 550
Db        547 CGAAGATGTATTCTACTCATCAACAAGGCAGAAAGTACC---CTCCCGCTTTCAGAAGGG 603

Qy        551 CysGlnThrLeuLeuSerProIleValSerCysGlyProProGlyValLeuLeuThrArg 570
Db        604 ACCCANACAGTATTGAGCCCCTCGGTGACCTGTGGACCCACAGGCCTCCTGCTGTGCCGC 663

Qy        571 ProValIleLeuAlaMetAspHisCysGlyGluProSerProAspSerTrpSerLeuArg 590
```

```
                  |||||||||||    |||    ||||||    |||    |||              |||        :::
Db        664 CCCGTCATCCTCACCATGCCCCACTGTGCCGAAGTCAGTGCCCGTGACTGGATCTTTCAG 723

Qy        591 LeuLysLysGlnSerCysGluGlySerTrpGluGlnAspValLeuHisLeuGlyGluGlu 610
              ||||||   |||:::   :::|||   |||        :::|||:::   |||   ||||||
Db        724 CTCAAGACCCAGGCCCACCAGGGCCACTGG---GANGAGGTGGTGACCCTGGATGAGGAG 780

Qy        611 AlaProSerHisLeuTyrTyrCysGlnLeuGluAlaSerAlaCysTyrValPheThrGlu 630
                    :::            ||||||||||||||       ||||||::::::      :::
Db        781 ACCCTGAACACACCCTGCTACTGCCAGCTGGAGCCCAGGGCCTGTCACATCCTGCTGGAC 840

Qy        631 GlnLeuGlyArgPhe-AlaLeuValGlyGluAlaLeuSerValAlaAlaAlaLysArgLe 650
              |||||||||   :::              ||||||:::    |||     :::|||   ||||||||
Db        841 CAGCTGGGCACCTACCGTGTTCACGGGCGAGTCCTATTCCCGCTCAGCAGTCAAGCGGCT 900

Qy        650 uLysLeuLeuLeuPheAlaProValAlaCysThrSerLeuGluTyrAsnIleArgValTy 670
              |:::|||   :::|||||||||         ||||||||||||||||||||:::::::||||||||
Db        901 CCAGCTGGCCGTCTTCGCCCCCGCCCTCTGCACCTCCCTGGAGTACAGCCTCCGGGTCTA 960

Qy        670 rCysLeuHisAspThrHisAspAlaLeuLysGluValValGlnLeuGluLysGlnLeuGl 690
              |||||||   ||||||         ||||||||||||||||:::::::||||||:::    |||||
Db        961 CTGCCTGGAGGACACGCCTGTAGCACTGAAGGAGGTGCTGGAGCTGGAGCGGACTCTGGG 1020

Qy        690 yGlyGlnLeuIleGlnGluProArgValLeuHisPheLysAspSerTyrHisAsnLeuAr 710
              ||||   |||:::::::||||||:::    |||   ||||||||||||||||||||||||||
Db       1021 CGGATACTTGGTGGAGGAGCCGAAACCGCTAATGTTCAAGGACAGTTACCACAACCTGCG 1080

Qy        710 gLeuSerIleHisAspValProSerSerLeuTrpLysSerLysLeuLeuValSerTyrGl 730
              ||||||:::||||||:::|||          |||:::||||||||||||||      |||||
Db       1081 CCTCTCCCTCCATGACCTCCCCCATGCCCATTGGAGGAGCAAGCTGCTGGCCAAATACCA 1140

Qy        730 nGluIleProPheTyrHisIleTrpAsnGlyThrGlnArgTyrLeuHisCysThrPheTh 750
              ||||||||||||||||||||||||:::|||:::|||:::    ||||||||||||||||||
Db       1141 GGAGATCCCCTTCTATCACATTTGGAGTGGCAGCCAGAAGGCCCTCCACTGCACTTTCAC 1200

Qy        750 rLeuGluArg-ValSerProSerThrSerAspLeuAlaCysLysLeuTrpValTrpGlnV 770
              ||||||||||             ::::::::::::|||   ||||||:::   |||   ||||
Db       1201 CCTGGAGAGGCACAGGCTTGGCCTCCACAGAGCTCACCTGCAAGATCTGCGTGCGGCAAG 1260

Qy        770 alGluGlyAspGlyGlnSerPheSerIleAsnPheAsnIleThrLysAspThrArgPheA 790
              |||||||:::||||||    |||   ::::::      :::       :::|||
Db       1261 TGGAAGGGGAGGGCCAGATATTCCAGCTGCATACCACTCTGGCA---GAGACACCTGCTG 1317

Qy        790 laGluLeuLeuAlaLeuGluSerGluAlaGly------ValProAlaLeuValGlyProS 808
                 |||      |||   |||      |||               :::||||||
Db       1318 GCTCCCTGGACACTCTCTGCTCTGCCCCTGGCAGCACTGTCACCACCCAGCTGGGACCTT 1377

Qy        808 erAlaPheLysIleProPheLeuIleArgGlnLysIleIleSerSerLeuAspProProC 828
                 |||||||||||||||      ||||||||||||||      :::|||||||||   |||
Db       1378 ATGCCTTCAAGATCCCACTGTCCATCCGCCAGAAGATATGCAACAGCCTAGATGCCCCCA 1437

Qy        828 ysArgArgGlyAlaAspTrpArgThrLeuAlaGlnLysLeuHisLeuAspSerHisLeuS 848
                    ||||||   |||||||||   |||||||||||||||     :::|||    :::|||:
Db       1438 ACTCACGGGGCAATGACTGGCGGATGTTAGCACAGAAGCTCTCTATGGACCGGTACCTGA 1497

Qy        848 erPhePheAlaSerLysProSerProThrAlaMetIleLeuAsnLeuTrpGluAlaArgH 868
              ::::|||||||:::|||      |||||||||      :::||||||:::||||||||||||
Db       1498 ATTACTTTGCCACCAAAGCGAGCCCCACGGNTGTGATCCTGGACCTCTGGGAAGCTCTGC 1557

Qy        868 isPheProAsnGlyAsnLeuSerGlnLeuAlaAlaAlaVal 881
                 :::|||:::|||:::      |||      :::|||:::
Db       1558 AGCAGGACGATGGGGACCTCAACAGCCTGNCGAGTGCCTTG 1598

RESULT 9
US-08-253-155A-17/c
; Sequence 17, Application US/08253155A
; Patent No. 5691147
;  GENERAL INFORMATION:
;    APPLICANT:  Gyuris, Jeno
;    APPLICANT:  Draetta, Giulio
;    TITLE OF INVENTION:  CDK4 Binding Proteins
;    NUMBER OF SEQUENCES:  95
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD
;      STREET:  60 State Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII(text)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/253,155A
;      FILING DATE:  02-JUN-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Vincent, Matthew P.
;      REGISTRATION NUMBER:  36,709
;      REFERENCE/DOCKET NUMBER:  MII-028
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 227-7400
;      TELEFAX:  (617) 227-5941
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  771 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-253-155A-17

Alignment Scores:
Pred. No.:                 8.7e-49      Length:           771
Score:                     600.50       Matches:          131
Percent Similarity:        67.47%       Conservative:     37
Best Local Similarity:     52.61%       Mismatches:       77
Query Match:               12.54%       Indels:           6
DB:                        1            Gaps:             2

US-09-970-944-2 (1-899) x US-08-253-155A-17 (1-771)

Qy        541 ProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleValSer 560
              |||   ||||||                  |||         :::|||:::|||||||||   |||:::
Db        769 CCTTATGATGTGCCA---GATTATGCCTCTCCCGAGACAGTATTGAGCCCCTCGGTGACC 713

Qy        561 CysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGly 580
              |||||||||   |||:::||||||   ||||||||||||||   |||   ||||||
Db        712 TGTGGACCCACAGGCCTCCTGCTGTGCCGCCCCGTCATCCTCACCATGCCCCACTGTGCC 653

Qy        581 GluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySerTrp 600
              |||   |||         |||         :::||||||   |||:::   :::|||   |||
Db        652 GAAGTCAGTGCCCGTGACTGGATCTTTCAGCTCAAGACCCAGGCCCACCAGGGCCACTGG 593

Qy        601 GluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGlnLeu 620
              |||   :::|||:::   |||   ||||||         :::         ||||||||||||
Db        592 GAG---GAGGTGGTGACCCTGGATGAGGAGACCCTGAACACACCCTGCTACTGCCAGCTG 536

Qy        621 GluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGlu 640
              |||      ||||||::::::      :::|||||||||   :::         ||||||
Db        535 GAGCCCAGGGCCTGTCACATCCTGCTGGACCAGCTGGGCACCTACGTGTTCACGGGCGAG 476

Qy        641 AlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAlaCys 660
              :::   |||   :::|||   ||||||   :::|||   :::|||||||||      |||
Db        475 TCCTATTCCCGCTCAGCAGTCAAGCGGTCCAGCTGGCCGTCTTCGCCCCCGCCCTCTGC 416

Qy        661 ThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuLys 680
              ||||||||||||||||   :::|||||||||||||||   |||         |||||||||
Db        415 ACCTCCCTGGAGTACAGACTCCGGGTCTACTGTCTGGAGGACACAGGCCTGTAGCACTGAAG 356

Qy        681 GluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgValLeu 700
              ||||||::::::||||||:::   |||||||||   |||::::::||||||:::   |||
Db        355 GAGGTGCTGGAGCTGGAGCGGACTCTGGGCGGATACTTGGTGGAGGAGCCGAAACCGCTA 296
```

```
Qy        701 HisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSerLeu 720
                  |||||||||||   |||||||||||    |||   |||||:::|||    :::
Db        295 ATGTTCAAGGACAGTTAACACAACCTGCGCTCTTCCTC-CATGACCTCCCCCATGCCCAT 237

Qy        721 TrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsnGly 740
               |||:::||||||||||||          |||||||||||||||||||||||||||:::|||
Db        236 TGGAGGAGCAAGCTGCTGGCCAAATACCAGGAGATCCCCTTCTATCACATTTGGAGTGGA 177

Qy        741 ThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSerAsp 760
                 :::|||    |||||||||||||||||||||||||
Db        176 GCCAGAAGGC--CTCCACTGCACTTTCACCCTGGAGAGGACAGGTTGGCCTCCACAGAGC 119

Qy        761 LeuAlaCysLys-LeuTrpValTrpGlnVal-GluGlyAspGlyGlnSerPheSerIleA 780
                     |||||| :::     |||     :::|||  |||||:::||||||    |||    ::::
Db        118 TCAACCTGCAAGGATCTGCGTGCGGGAAGTTGGAAGGGGAGGGACAGATATTCCAGCTGC 59

Qy        780 snPheAsnIleThrLysAspThr 787
               ::      :::    :::||||||
Db         58 ATACCACTCTGGGAGAGGACACC 36
```

RESULT 10
US-09-262-537-48
; Sequence 48, Application US/09262537
; Patent No. 6479256
; GENERAL INFORMATION:
;  APPLICANT: Hayflick, Joel
;  TITLE OF INVENTION: Lectomedin Materials and Methods
;  FILE REFERENCE: 27866/35307
;  CURRENT APPLICATION NUMBER: US/09/262,537
;  CURRENT FILING DATE: 1999-03-04
;  EARLIER APPLICATION NUMBER: 60/076,782
;  EARLIER FILING DATE: 1998-03-04
;  NUMBER OF SEQ ID NOS: 64
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 48
;   LENGTH: 5749
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-262-537-48

```
Alignment Scores:
Pred. No.:                 7.47e-15       Length:          5749
Score:                     267.50         Matches:         108
Percent Similarity:        38.69%         Conservative:    58
Best Local Similarity:     25.17%         Mismatches:      132
Query Match:               5.59%          Indels:          134
DB:                        4              Gaps:            21
```

US-09-970-944-2 (1-899) x US-09-262-537-48 (1-5749)

```
Qy         30 ThrValAlaAsnProValProGly------------------------------------ 37
              |||:::     ||||||   |||
Db        901 ACGCTGGAAAATCCAGATCCAACCAAATATAGCATTACCTGAAATTTTCCAAAAAGGAC 960

Qy         38 ---------AlaAsnProAspLeuLeu------------ProHisPheLeuValGluPro 50
                       :::|||     ||||||              ||||||    |||
Db        961 CTTAGCTGCTCTAACTTTTCACTCCTGGCTTATCAGTTTGATCATTTTTCCCATGAAAAA 1020

Qy         51 GluAspValTyrIleValLysAsnLysProValLeuLeuValCysLysAlaValProAla 70
                             :::    ||||||        ::::::    :::|||     :::         |||
Db       1021 ATAAAGGATCTTTTAAGAAAGAATCATTCTATAATGCAACTCTGCAATTCCAAGAATGCT 1080

Qy         71 ThrGlnIlePhePheLysCysAsnGlyGluTrpValArgGlnValAspHisValIleGlu 90
                   :::|||   :::    :::       ::::::    |||:::       |||
Db       1081 TTC---GTTTTCTACAGTATGATAAAAATTTTATT---CAAATACGTCGAGTATTT--- 1131

Qy         91 ArgSerThrAspGlySerSerGlyGluProThrMetGluValArgIleAsnValSerArg 110
                                                 ||||||        ||||||           |||
Db       1132 ---------------------------CCAACTAATTT-CCCAGGATTACAGAAAAAAGG 1163

Qy        111 GlnGlnValGluLys------------ValPheGlyLeuGluGlu--------------- 121
                 :::      :::              |||||||||:::|||:::
Db       1164 GGAAGAAGATCAGAAATCTTTTTTGAGTTTTGGTATTGAACAAGGTCAGCCCAAGCCA 1223

Qy        122 TyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAlaTyr 141
                 |||         |||:::     :::             |||     :::|||
Db       1224 GTTTGGTTGCCATGTATTATGTACTTGGTTGGAGAGCTGCTTAAAATCAGAAAATGGGAG 1283

Qy        142 IleArgIle---------------------AlaArgLeuArgLysAsnPhe--------- 151
                 ||||||                       |||         |||
Db       1284 AACAGAATCATGTGGGATCATGTATACAAAATGCACCTGCCCTCAGCATTTGGGAGAGTG 1343

Qy        152 ---GluGlnGluProLeuAlaLysGluValSer-----------------------Leu 162
                 ::::::     |||:::|||      |||
Db       1344 GGGGATCGACGACCAGTCGCTGATTTTGTTAAATAACGTGGTGTTACCCCTGAATGAGCA 1403

Qy        163 GluGlnGlyIleValLeuPro------CysArgProProGlu---GlyIleProProAla 179
              ::::::|||:::          |||        |||:::||||||:::       |||   |||
Db       1404 GACAGAGGGCTGCCTGA-CCCAGGAGCTGCAAACCACCCAAGTCTGCAATCTTACCAGGG 1462

Qy        180 GluValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspPro------Asn 197
                                                ||||||      |||      |||
Db       1463 AGG---------------------------------CCAAGCGACCACCCAAAGAAGAAT 1489

Qy        198 ValTyrIleThrArgGluHisSerLeuValValArgGlnAlaArg--------------- 212
              :::    :::          :::|||::::::       ::::::    |||
Db       1490 TTGGA-ATGATGGGAGATCATACAATTAAAAGTCAGCGACCTCGATCTGTTCATGAAAAA 1548

Qy        213 ---------------LeuAlaAspThrAlaAsnTyr----------------------- 219
                             ||||||    |||     :::
Db       1549 AGGGTCCCTCAGGAACAAGCTGATGCTGCTAAATTTATGGCACAAACTGGTGAATCTGGT 1608

Qy        219 ------------------------------------------------------------ 219

Db       1609 GTGGAAGAGTGGTCCCAGTGGAGCACATGTTCGGTTACTTGTGGTCAAGGGTCGCAGGTG 1668

Qy        220 ---------ThrCysValAla-----------------LysAsnIleValAlaArgArg 230
                       |||||||||:::                            :::
Db       1669 CGAACCAGAACTTGTGTATCACCTTACGGGACACACTGCAGCGGCCCATTAAGAGAATCA 1728

Qy        231 ArgSerAlaSerAlaAlaValIleValTyrValAsnGlyGlyTrpSerThrTrpThrGlu 250
              |||          :::          :::      |||:::|||    |||        |||:::
Db       1729 AGGGTTTGCAATAACACTGCCCTCTGTCCAGTACACGGAGTATGGGAGGAATGGTCACCA 1788

Qy        251 TrpSerValCysSerAlaSerCysGlyArgGlyTrpGlnLysArgSerArgSerCysThr 270
              ||||||:::||||||     :::||||||||||||     :::     |||:::||||||||||||
Db       1789 TGGAGTTTATGTTCATTTACATGTGGTCGAGGCCAAAGAACAAGAACAAGGTCATGCACA 1848

Qy        271 AsnProAlaProLeuAsnGlyGlyAlaPheCysGluGlyGlnAsnValHisAspArgThr 290
                       |||      ||||||      |||||||||             |||     :::
Db       1849 ------CCTCCTCAGTATGGAGGAAGGCCGTGTGAAGGACCTGAAACACATCATAAGCCT 1902

Qy        291 ValSerSerLeuLeuValSerValAspGlySerTrpSerProTrpSerLysTrpSerAla 310
                 :::      |||        |||||||||     |||         ||||||    ||||||
Db       1903 TGTAATATTGCTCTTTGCCCAGTTGATGGACAGTGGCAAGAGTGGAGTTCGTGGAGCCAG 1962

Qy        311 CysGlyLeuAspCys---------ThrHisTrpArgSerArgGluCysSerAspProAla 327
              |||    :::     |||                |||      |||||||||:::|||:::        |||
Db       1963 TGCTCAGTAACGTGCTCGAATGGGACTCAGCAGAGAAGCCGGCAGTGCACT------GCA 2016

Qy        328 ProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeuAspThrArgAsnCysThrSer 347
                    :::||||||     ||||||:::|||              ::::::|||     |||     :::
Db       2017 GCTGCCCATGGAGGCTCCGAATGCAGAGGGCCATGGGCAGAAAGCAGAGAGTGCTATAAC 2076

Qy        348 AspLeuCysValHisSerAlaSerGly 356
                    |||      :::|||:::|||
Db       2077 CCTGAATGT------ACAGCCAATGGT 2097
```

RESULT 11
US-08-808-982-4
; Sequence 4, Application US/08808982
; Patent No. 5939271
;  GENERAL INFORMATION:
;    APPLICANT:  Tessier-Lavigne, Marc
;    APPLICANT:  Leonardo, E. David
;    APPLICANT:  Hink, Lindsay

```
;    APPLICANT:  Masu, Masayuki
;    APPLICANT:  Kazuko, Keino-Masu
;    TITLE OF INVENTION:  Netrin Receptors
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SCIENCE & TECHNOLOGY LAW GROUP
;      STREET:  268 BUSH STREET, SUITE 3200
;      CITY:  SAN FRANCISCO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/808,982
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  OSMAN, RICHARD A
;      REGISTRATION NUMBER:  36,627
;      REFERENCE/DOCKET NUMBER:  UC96-217
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 343-4341
;      TELEFAX:  (415) 343-4342
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  305 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-808-982-4

Alignment Scores:
Pred. No.:              1.08e-15     Length:        305
Score:                  253.00       Matches:       56
Percent Similarity:     71.43%       Conservative:  14
Best Local Similarity:  57.14%       Mismatches:    28
Query Match:            5.29%        Indels:        2
DB:                     2            Gaps:          0

US-09-970-944-2 (1-899) x US-08-808-982-4 (1-305)

Qy        612 ProSerHisLeuTyrTyrCysGlnLeuGluAlaSerAlaCysTyrValPheThrGluGln 631
              |||   |||             ||||||||    |||||   :::       :::|||
Db         14 CCTGAACACACCCTGCTACTGCAGCTGGAGCCCAGGGCCTGTAC-ATCCTGCTGGACCAG 72

Qy        632 LeuGlyArgPheAlaLeuValGlyGluAlaLeuSerValAlaAlaAlaLysArgLeuLys 651
              |||||    :::        ||||||:::   |||    :::|||   |||||||||:::
Db         73 CTGGGCACCTACGTTTTCACGGGCGAGTCCTATTCCCGCTCAGCAGTCAAGCGGCTCCAG 132

Qy        652 LeuLeuLeuPheAlaProValAlaCysThrSerLeuGluTyrAsnIleArgValTyrCys 671
              |||      ||||||||      |||||||||||||||||:::::::||||||||||||
Db        133 CTGGCCGT-TTCGCCCCCGCCCTCTGCACCTCCCTGGAGTACAGCCTCCGGGTCTACTGC 191

Qy        672 LeuHisAspThrHisAspAlaLeuLysGluValValGlnLeuGluLysGlnLeuGlyGly 691
              |||   ||||||      ||||||||||||||:::::::||||||:::      |||||||||
Db        192 CTGGAGGACACGCCTGTAGCACTGAAGGAGGTGCTGGAGCTGGAGCGGACTCTGGGCGGA 251

Qy        692 GlnLeuIleGlnGluProArgValLeuHisPheLysAspSerTyrHisAsnLeu 709
                 |||::::::||||||:::   |||   ||||||||||||||||||||||||
Db        252 TACTTGGTGGAGGAGCCGAAACCGCTAATGTTCAAGGACAGTTACCACAACCTT 305
```

RESULT 12
US-09-306-902A-4
; Sequence 4, Application US/09306902A
; Patent No. 6277585
;   GENERAL INFORMATION:
;        APPLICANT: Tessier-Lavigne, Marc
;                   Leonardo, E. David
;                   Hink, Lindsay
;                   Masu, Masayuki
;                   Kazuko, Keino-Masu

```
;       TITLE OF INVENTION: Netrin Receptors
;       NUMBER OF SEQUENCES: 9
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
;            STREET: 268 BUSH STREET, SUITE 3200
;            CITY: SAN FRANCISCO
;            STATE: CALIFORNIA
;            COUNTRY: USA
;            ZIP: 94104
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/306,902A
;            FILING DATE: 07-May-1999
;            CLASSIFICATION: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: OSMAN, RICHARD A
;            REGISTRATION NUMBER: 36,627
;            REFERENCE/DOCKET NUMBER: UC96-217
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (415) 343-4341
;            TELEFAX: (415) 343-4342
;  INFORMATION FOR SEQ ID NO: 4:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 305 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: double
;            TOPOLOGY: linear
;       MOLECULE TYPE: cDNA
;       SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-306-902A-4

Alignment Scores:
Pred. No.:              1.08e-15     Length:        305
Score:                  253.00       Matches:       56
Percent Similarity:     71.43%       Conservative:  14
Best Local Similarity:  57.14%       Mismatches:    28
Query Match:            5.29%        Indels:        2
DB:                     3            Gaps:          0

US-09-970-944-2 (1-899) x US-09-306-902A-4 (1-305)

Qy        612 ProSerHisLeuTyrTyrCysGlnLeuGluAlaSerAlaCysTyrValPheThrGluGln 631
              |||   |||             ||||||||    |||||   :::       :::|||
Db         14 CCTGAACACACCCTGCTACTGCAGCTGGAGCCCAGGGCCTGTAC-ATCCTGCTGGACCAG 72

Qy        632 LeuGlyArgPheAlaLeuValGlyGluAlaLeuSerValAlaAlaAlaLysArgLeuLys 651
              |||||    :::        ||||||:::   |||    :::|||   |||||||||:::
Db         73 CTGGGCACCTACGTTTTCACGGGCGAGTCCTATTCCCGCTCAGCAGTCAAGCGGCTCCAG 132

Qy        652 LeuLeuLeuPheAlaProValAlaCysThrSerLeuGluTyrAsnIleArgValTyrCys 671
              |||      ||||||||      |||||||||||||||||:::::::||||||||||||
Db        133 CTGGCCGT-TTCGCCCCCGCCCTCTGCACCTCCCTGGAGTACAGCCTCCGGGTCTACTGC 191

Qy        672 LeuHisAspThrHisAspAlaLeuLysGluValValGlnLeuGluLysGlnLeuGlyGly 691
              |||   ||||||      ||||||||||||||:::::::||||||:::      |||||||||
Db        192 CTGGAGGACACGCCTGTAGCACTGAAGGAGGTGCTGGAGCTGGAGCGGACTCTGGGCGGA 251

Qy        692 GlnLeuIleGlnGluProArgValLeuHisPheLysAspSerTyrHisAsnLeu 709
                 |||::::::||||||:::   |||   ||||||||||||||||||||||||
Db        252 TACTTGGTGGAGGAGCCGAAACCGCTAATGTTCAAGGACAGTTACCACAACCTT 305
```

RESULT 13
US-08-985-526-4
; Sequence 4, Application US/08985526
; Patent No. 6080728
;   GENERAL INFORMATION:

```
;    APPLICANT:  Mixson, James A
;    TITLE OF INVENTION:  CARRIER:DNA COMPLEXES CONTAINING DNA
;    TITLE OF INVENTION:  ENCODING ANTI-ANGIOGENIC PEPTIDES AND THEIR USE IN GENE
;    TITLE OF INVENTION:  THERAPY
;    NUMBER OF SEQUENCES:  43
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Connolly, Bove, Lodge, & Hutz
;      STREET:  1220 Market Street, P.O. Box 2207
;      CITY:  Wilmington
;      STATE:  Delaware
;      COUNTRY:  U.S.A.
;      ZIP:  19899
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/985,526
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/608,845
;      FILING DATE:  16-JUL-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McMorrow Jr., Robert G
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (302) 658-9141
;      TELEFAX:  (302) 658-5613
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1326 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-985-526-4

Alignment Scores:
Pred. No.:              2.82e-14      Length:        1326
Score:                  250.00        Matches:       104
Percent Similarity:     35.81%        Conservative:  50
Best Local Similarity:  24.19%        Mismatches:    152
Query Match:            5.22%         Indels:        126
DB:                     3             Gaps:          17

US-09-970-944-2 (1-899) x US-08-985-526-4 (1-1326)

Qy         13 GlyIleValLeuAlaAlaTrpLeuArgGlySerGlyAlaGlnGlnSerAlaThrValAla 32
Db         92 GGACTGTTGATAGCTGCA-----------------------CTGAGTGTCACTGTCAGA 127

Qy         33 AsnProValProGlyAlaAsn--------------------ProAspLeuLeuProHis 45
Db        128 ACTCAGTTACCATCTGCAAAAAGGTGTCCTGCCCCATCATGCCCTGCTCCAATGCCACAG 187

Qy         46 PheLeuValGlu------------------------------------------------ 49
Db        188 TTCCTGATGGAGAATGCTGTCCTCGCTGTTGGCCCAGCGACTCTGCGGACGATGGCTGGT 247

Qy         50 ---------ProGluAspValTyrIleValLysAsnLysProValLeuLeuValCysLys 66
Db        248 CTCCATGGTCCGAGTGGACCTCCTGTTCTACGAGCTGTGGCAATGGAATTCAGCAGCGCG 307

Qy         67 AlaValProAlaThrGlnIlePhePheLysCysAsnGlyGluTrpValArg---GlnVal 85
Db        308 GCCGCTCCTGCGAT-AGCCTCAACAACCGATGTGAGGGCTCCTCGGTCCAGACACGGACC 366

Qy         86 AspHisValIleGluArgSerThrAspGlySerSerGlyGluProThrMetGluValArg 105
Db        367 TGCCACATT--------------------------------------------------- 375

Qy        106 IleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGluGluTyrTrpCysGln 125
Db        376 ---------------CAGGAGTGTGACAAAAGATTTAAACAGGATGGTGGCTGGAGCCAC 420

Qy        126 CysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAlaTyrIleArgIleAla 145
Db        421 TGGTCCCCGTGGTCATCTTGTTCTGTGACATGTGGTGATGGTGTGATCACAAGGATC--- 477

Qy        146 ArgLeuArgLysAsnPheGluGlnGluPro-LeuAlaLysGluValSerLeuGluGlnGl 165
Db        478 CGGCTCTGCAACTCTCCCAGCCCCCAGATGAATGGGAAACCCTGTGAAGGCGAAGCGCGG 537

Qy        165 yIleValLeuProCysArgProProGluGly----------------------------- 175
Db        538 GAGACCAAAGCCTGCAAGAAAGACGCCTGCCCCATCAATGGAGGCTGGGGTCCTTGGTCA 597

Qy        176 ----------------IleProProAlaGluValGluTrpLeuArgAsnGluAspLeuVa 190
Db        598 CCATGGGACATCTGTTCTGTCACCTGTGGAG---GAGGGGTACAGAAACGTAGTCGTCTC 654

Qy        190 lAspProSerLeuAsp-------------------------------------------- 195
Db        655 TGCGTCGACTCTAGAATGACTGAAGAGAACAAAGAGTTGGCCAATGAGCTGAGGCGGCCT 714

Qy        196 -ProAsnValTyrIleThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAs 215
Db        715 CCCCTATGCT---ATCACAACGGAGTTCAGTACAGAAATAACGAGGAATGGACTGTT-GA 770

Qy        215 pThrAlaAsnTyrThrCysValAlaLysAsnIleValAlaArgArgArgSerAlaSerAl 235
Db        771 TAGCTGCACT---GAGTGTCACTGTCAGAACTCAGTTACCATCTGCAAAAAGGTGTCCTG 827

Qy        235 aAlaValIleValTyrValAsnGly----------------------------------- 243
Db        828 CCCCATCATGCCCTGCTCCAATGCCACAGTTCCTGATGGAGAATGCTGTCCTCGCTGTTG 887

Qy        244 ---------------------GlyTrpSerThrTrpThrGluTrpSerValCysSerAl 256
Db        888 GCCCAGCGACTCTGCGGACGATGGCTGGTCTCCATGGTCCGAGTGGACCTCCTGTTCTAC 947

Qy        256 aSerCysGlyArgGlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAs 276
Db        948 GAGCTGTGGCAATGGAATTCAGCAGCGCGGCCGCTCCTGCGATAGC---------CTCAA 998

Qy        276 nGlyGlyAlaPheCysGluGlyGlnAsnValHisAspArgThrVal-----------Se 292
Db        999 CAACCGA------TGTGAGGGCTCCTCGGTCCAGACACGGACCTGCCACATTCAGGAGTG 1052

Qy        292 rSerLeuLeuValSerValAspGlySerTrpSerProTrpSerLysTrpSerAlaCysGl 312
Db       1053 TGACAAAAGATTTAAACAGGATGGTGGCTGGAGCCACTGGTCCCCGTGGTCATCTTGTTC 1112

Qy        312 yLeuAspCys---------ThrHisTrpArgSerArgGluCysSerAspProAlaProAr 329
Db       1113 TGTGACATGTGGTGATGGTGTGATCACAAGGATCCGGCTCTGCAACTCTCCCAGCCCCCA 1172

Qy        329 gAsnGlyGlyGluGluCysGlnGlyThrAspLeuAspThrArgAsnCysThrSerAspLe 349
Db       1173 GATGAATGGGAAACCCTGTGAAGGCGAAGCGCGGGAGACCAAAGCCTGCAAGAAAGACGC 1232

Qy        349 uCys---ValHisSerAlaSerGlyPro 357
Db       1233 CTGCCCCATCAATGGAGGCTGGGGTCCT 1260

RESULT 14
US-09-919-172-64
; Sequence 64, Application US/09919172
; Patent No. 6673545
; GENERAL INFORMATION:
;  APPLICANT: Faris, Mary
;  APPLICANT:  Turner, Christopher M.
;  TITLE OF INVENTION: PROSTATE CANCER MARKERS
;  FILE REFERENCE: PA-0036 US
;  CURRENT APPLICATION NUMBER: US/09/919,172
;  CURRENT FILING DATE:  2001-07-30
```

```
;  PRIOR APPLICATION NUMBER: 60/222,469
;  PRIOR FILING DATE: 2000-07-28
;  NUMBER OF SEQ ID NOS: 102
;  SOFTWARE: PERL Program
; SEQ ID NO 64
;   LENGTH: 7231
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. 6673545 1399366.20
;   NAME/KEY: unsure
;   LOCATION: 5601, 5609, 7107
;   OTHER INFORMATION: a, t, c, g, or other
US-09-919-172-64

Alignment Scores:
Pred. No.:              5.64e-13      Length:        7231
Score:                  250.00        Matches:       48
Percent Similarity:     45.81%        Conservative:  23
Best Local Similarity:  30.97%        Mismatches:    66
Query Match:            5.22%         Indels:        18
DB:                     4             Gaps:          3

US-09-970-944-2 (1-899) x US-09-919-172-64 (1-7231)

Qy       207 ValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThrCysValAlaLysAsnIle 226
             :::  ::::::  |||      ||||||  |||     |||            ::::::
Db      1364 ATTCAGCAGCGCGGCCGCTCCTGCGATACGCTCAACAACCGATGTGAGGGCTCCTCGGTC 1423

Qy       227 ValAlaArg------------------------ArgArgSerAlaSerAlaAlaValIle 238
                |||                          :::|||
Db      1424 CAGACACGGACCTGCCACATTCAGGAGTGTGACAAGAGATTTAAACAG------------ 1471

Qy       239 ValTyrValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCys 258
                      :::||||||||||||   |||:::   ||||||  ||||||   :::|||
Db      1472 ---------GATGGTGGCTGGAGCCACTGGTCCCCGTGGTCATCTTGTTCTGTGACATGT 1522

Qy       259 GlyArgGlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGly 278
             |||   |||          |||   |||   |||   :::|||:::|||       |||
Db      1523 GGTGATGGTGTGATCACAAGGATCCGGCTCTGCAACTCTCCCAGCCCCCAGATGAACGGG 1582

Qy       279 AlaPheCysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerVal 298
                   |||||||||:::               :::                   :::
Db      1583 AAACCCTGTGAAGGCGAAGCGCGGGAGACCAAAGCCTGCAAGAAAGACGCCTGCCCCATC 1642

Qy       299 AspGlySerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCys--------- 315
             :::|||   |||   |||||||||   |||      |||   :::   |||
Db      1643 AATGGAGGCTGGGGTCCTTGGTCACCATGGGACATCTGTTCTGTCACCTGTGGAGGAGGG 1702

Qy       316 ThrHisTrpArgSerArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCys 335
                      |||||||||   |||::::::|||   |||:::   ||||||::::::|||
Db      1703 GTACAGAAACGTAGTCGTCTCTGCAACAACCCCACACCCCAGTTTGGAGGCAAGGACTGC 1762

Qy       336 GlnGlyThrAspLeuAspThrArgAsnCysThrSerAspLeuCys 350
                |||         :::   :::   |||             |||
Db      1763 GTTGGTGATGTAACAGAAAACCAGATCTGCAACAAGCAGGACTGT 1807

RESULT 15
US-08-985-526-2
; Sequence 2, Application US/08985526
; Patent No. 6080728
;  GENERAL INFORMATION:
;    APPLICANT:  Mixson, James A
;    TITLE OF INVENTION:  CARRIER:DNA COMPLEXES CONTAINING DNA
;    TITLE OF INVENTION:  ENCODING ANTI-ANGIOGENIC PEPTIDES AND THEIR USE IN GENE
;    TITLE OF INVENTION:  THERAPY
;    NUMBER OF SEQUENCES:  43
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Connolly, Bove, Lodge, & Hutz
;      STREET:  1220 Market Street, P.O. Box 2207
;      CITY:  Wilmington
;      STATE:  Delaware
;      COUNTRY:  U.S.A.
;      ZIP:  19899
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/985,526
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/608,845
;      FILING DATE:  16-JUL-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McMorrow Jr., Robert G
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (302) 658-9141
;      TELEFAX:  (302) 658-5613
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  657 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-985-526-2

Alignment Scores:
Pred. No.:              2.24e-14      Length:        657
Score:                  245.50        Matches:       49
Percent Similarity:     53.28%        Conservative:  16
Best Local Similarity:  40.16%        Mismatches:    44
Query Match:            5.13%         Indels:        13
DB:                     3             Gaps:          5

US-09-970-944-2 (1-899) x US-08-985-526-2 (1-657)

Qy       244 GlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArgGlyTrpGln 263
             |||||||||      |||:::|||||||:::   ||||||   |||||||||   |||   |||
Db       241 GGCTGGTCTCCATGGTCCGAGTGGACCTCCTGTTCTACGAGCTGTGGCAATGGAATTCAG 300

Qy       264 LysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPheCysGluGly 283
             :::|||   ||||||||||   :::            ||||||          ||||||||
Db       301 CAGCGCGGCCGCTCCTGCGATAGC---------CTCAACAACCGA------TGTGAGGGC 345

Qy       284 GlnAsnValHisAspArgThrVal------------SerSerLeuLeuValSerValAsp 299
                :::|||      ||||||                                    |||
Db       346 TCCTCGGTCCAGACACGGACCTGCCACATTCAGGAGTGTGACAAAAGATTTAAACAGGAT 405

Qy       300 GlySerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCys---------Thr 316
             |||   ||||||   ||||||   ||||||:::|||      :::   |||
Db       406 GGTGGCTGGAGCCACTGGTCCCCGTGGTCATCTTGTTCTGTGACATGTGGTGATGGTGTG 465

Qy       317 HisTrpArgSerArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGln 336
                   |||   |||   |||:::   |||:::|||:::         |||:::   |||:::
Db       466 ATCACAAGGATCCGGCTCTGCAACTCTCCCAGCCCCCAGATGAATGGGAAACCCTGTGAA 525

Qy       337 GlyThrAspLeuAspThrArgAsnCysThrSerAspLeuCys---ValHisSerAlaSer 355
             |||         :::|||:::   |||         |||   |||    ::::::
Db       526 GGCGAAGCGCGGGAGACCAAAGCCTGCAAGAAAGACGCCTGCCCCATCAATGGAGGCTGG 585

Qy       356 GlyPro 357
             ||||||
Db       586 GGTCCT 591

Search completed: October  6, 2004, 22:39:13
Job time : 210 secs
```

```
OM protein - nucleic search, using frame_plus_p2n model

Run on:         October  6, 2004, 18:22:36 ; Search time 1039 Seconds
                                             (without alignments)
                                             4386.468 Million cell updates/sec

Title:          US-09-970-944-2
Perfect score:  4787
Sequence:       1 MAVRPGLWPALLGIVLAAWL..........AVAGLGQPDAGLFTVSEAEC 899

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       3340653 seqs, 2534783454 residues

Total number of hits satisfying chosen parameters:      6681306

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09970944/runat_05102004_112011_1685/app_query.fasta_1.1095
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09970944_@CGN_1_1_622_@runat_05102004_112011_1685
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      Published_Applications_NA:*
                1:  /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
                2:  /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
                3:  /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
                4:  /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
                5:  /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
                6:  /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
                7:  /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
                8:  /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
                9:  /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
                10:  /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
                11:  /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
                12:  /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
                13:  /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
                14:  /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
                15:  /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
                16:  /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
                17:  /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
                18:  /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
                19:  /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 4787 | 100.0 | 2881 | 11 | US-09-970-944-1 | Sequence 1, Appli |
| 2 | 4698.5 | 98.2 | 2752 | 13 | US-09-918-779-1 | Sequence 1, Appli |
| 3 | 4698.5 | 98.2 | 2752 | 17 | US-10-624-932-1 | Sequence 1, Appli |
| 4 | 4545.5 | 95.0 | 2697 | 15 | US-10-240-154-15 | Sequence 15, Appl |
| 5 | 4545.5 | 95.0 | 3014 | 10 | US-09-933-261-1 | Sequence 1, Appli |
| 6 | 4545.5 | 95.0 | 3014 | 15 | US-10-256-702-1 | Sequence 1, Appli |
| 7 | 4379.5 | 91.5 | 3580 | 17 | US-10-311-623-13 | Sequence 13, Appl |
| 8 | 2694 | 56.3 | 1787 | 10 | US-09-933-261-2 | Sequence 2, Appli |
| 9 | 2694 | 56.3 | 1787 | 15 | US-10-256-702-2 | Sequence 2, Appli |
| 10 | 2490.5 | 52.0 | 2995 | 13 | US-09-972-211-55 | Sequence 55, Appl |
| 11 | 2490.5 | 52.0 | 2995 | 13 | US-10-096-625-55 | Sequence 55, Appl |
| 12 | 2484 | 51.9 | 2860 | 13 | US-10-087-684-1 | Sequence 1, Appli |
| 13 | 2484 | 51.9 | 2860 | 13 | US-10-087-684-3 | Sequence 3, Appli |
| 14 | 2484 | 51.9 | 2860 | 13 | US-10-218-779-1 | Sequence 1, Appli |
| 15 | 2484 | 51.9 | 2860 | 13 | US-10-218-779-3 | Sequence 3, Appli |
| 16 | 2479 | 51.8 | 3485 | 9 | US-09-816-828-18 | Sequence 18, Appl |
| 17 | 2479 | 51.8 | 3884 | 13 | US-10-147-493-145 | Sequence 145, App |
| 18 | 2479 | 51.8 | 3884 | 13 | US-10-145-127-145 | Sequence 145, App |
| 19 | 2479 | 51.8 | 3884 | 13 | US-10-160-503-145 | Sequence 145, App |
| 20 | 2479 | 51.8 | 3884 | 13 | US-10-143-118-145 | Sequence 145, App |
| 21 | 2479 | 51.8 | 3884 | 13 | US-10-144-993-145 | Sequence 145, App |
| 22 | 2479 | 51.8 | 3884 | 13 | US-10-158-787-145 | Sequence 145, App |
| 23 | 2479 | 51.8 | 3884 | 13 | US-10-140-024-145 | Sequence 145, App |
| 24 | 2479 | 51.8 | 3884 | 13 | US-10-140-808-145 | Sequence 145, App |
| 25 | 2479 | 51.8 | 3884 | 13 | US-10-152-405-145 | Sequence 145, App |
| 26 | 2479 | 51.8 | 3884 | 13 | US-10-127-852A-145 | Sequence 145, App |
| 27 | 2479 | 51.8 | 3884 | 13 | US-10-127-900A-145 | Sequence 145, App |
| 28 | 2479 | 51.8 | 3884 | 13 | US-10-128-685A-145 | Sequence 145, App |
| 29 | 2479 | 51.8 | 3884 | 13 | US-10-131-820A-145 | Sequence 145, App |
| 30 | 2479 | 51.8 | 3884 | 13 | US-10-142-886-145 | Sequence 145, App |
| 31 | 2479 | 51.8 | 3884 | 13 | US-10-146-728-145 | Sequence 145, App |
| 32 | 2479 | 51.8 | 3884 | 13 | US-10-146-786-145 | Sequence 145, App |
| 33 | 2479 | 51.8 | 3884 | 13 | US-10-147-499-145 | Sequence 145, App |
| 34 | 2479 | 51.8 | 3884 | 13 | US-10-157-798-145 | Sequence 145, App |
| 35 | 2479 | 51.8 | 3884 | 15 | US-10-028-072-145 | Sequence 145, App |
| 36 | 2479 | 51.8 | 3884 | 15 | US-10-121-049-145 | Sequence 145, App |
| 37 | 2479 | 51.8 | 3884 | 15 | US-10-123-904-145 | Sequence 145, App |
| 38 | 2479 | 51.8 | 3884 | 15 | US-10-140-470-145 | Sequence 145, App |
| 39 | 2479 | 51.8 | 3884 | 15 | US-10-175-746-145 | Sequence 145, App |
| 40 | 2479 | 51.8 | 3884 | 15 | US-10-176-918-145 | Sequence 145, App |
| 41 | 2479 | 51.8 | 3884 | 15 | US-10-176-921-145 | Sequence 145, App |
| 42 | 2479 | 51.8 | 3884 | 15 | US-10-137-865-145 | Sequence 145, App |
| 43 | 2479 | 51.8 | 3884 | 15 | US-10-140-474-145 | Sequence 145, App |
| 44 | 2479 | 51.8 | 3884 | 15 | US-10-142-431-145 | Sequence 145, App |
| 45 | 2479 | 51.8 | 3884 | 15 | US-10-143-114-145 | Sequence 145, App |

ALIGNMENTS

```
RESULT 1
US-09-970-944-1
; Sequence 1, Application US/09970944
; Publication No. US20030204052A1
; GENERAL INFORMATION:
;  APPLICANT: Herrman, John L
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Shimkets, Richard A
;  TITLE OF INVENTION: No. US20030204052A1el Proteins and Nucleic Acids Encoding Same and
;  TITLE OF INVENTION:  Antibodies Directed Against these Proteins
;  FILE REFERENCE: 21402-138
;  CURRENT APPLICATION NUMBER: US/09/970,944
;  CURRENT FILING DATE:  2002-05-02
;  PRIOR APPLICATION NUMBER: 60/237,862
;  PRIOR FILING DATE: 2000-10-04
;  NUMBER OF SEQ ID NOS: 62
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 2881
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-970-944-1
```

```
Alignment Scores:
Pred. No.:                0              Length:          2881
Score:                    4787.00        Matches:         899
Percent Similarity:       100.00%        Conservative:    0
Best Local Similarity:    100.00%        Mismatches:      0
Query Match:              100.00%        Indels:          0
DB:                       11             Gaps:            0

US-09-970-944-2 (1-899) x US-09-970-944-1 (1-2881)

Qy          1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         87 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 146

Qy         21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 206

Qy         41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 266

Qy         61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 326

Qy         81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCG 386

Qy        101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 446

Qy        121 GluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 506

Qy        141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 566

Qy        161 SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 626

Qy        181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 686

Qy        201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 746

Qy        221 CysValAlaLysAsnIleValAlaArgArgArgSerAlaSerAlaAlaValIleValTyr 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        747 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 806

Qy        241 ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        807 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 866

Qy        261 GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        867 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 926

Qy        281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        927 TGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGC 986

Qy        301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        987 AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC 1046

Qy        321 ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1047 CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG 1106

Qy        341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1107 GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG 1166

Qy        361 AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuLeuValLeu 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1167 GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC 1226

Qy        381 IleLeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIle 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1227 ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1286

Qy        401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1287 CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1346

Qy        421 LeuThrIleGlnProAspLeuSerThrThrThrThrTyrGlnGlySerLeuCysProArg 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1347 CTCACCATCCAGCCGGACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGG 1406

Qy        441 GlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeuGly 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1407 CAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGT 1466

Qy        461 GlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheValSer 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1467 GGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCC 1526

Qy        481 ArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThrTyr 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1527 CGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT 1586

Qy        501 GlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeuLeu 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1587 GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTC 1646

Qy        521 IleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHisLys 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1647 ATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG 1706

Qy        541 ProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleValSer 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1707 CCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC 1766

Qy        561 CysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGly 580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1767 TGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG 1826

Qy        581 GluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySerTrp 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1827 GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG 1886

Qy        601 GluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGlnLeu 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1887 GAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTG 1946

Qy        621 GluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGlu 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1947 GAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAG 2006

Qy        641 AlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAlaCys 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2007 GCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGC 2066

Qy        661 ThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuLys 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2067 ACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAG 2126
```

```
Qy        681 GluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgValLeu 700
Db       2127 GAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTG 2186

Qy        701 HisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSerLeu 720
Db       2187 CACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTG 2246

Qy        721 TrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsnGly 740
Db       2247 TGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGC 2306

Qy        741 ThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSerAsp 760
Db       2307 ACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGAC 2366

Qy        761 LeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIleAsn 780
Db       2367 CTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAAC 2426

Qy        781 PheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAlaGly 800
Db       2427 TTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGG 2486

Qy        801 ValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLysIle 820
Db       2487 GTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATA 2546

Qy        821 IleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGlnLys 840
Db       2547 ATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAA 2606

Qy        841 LeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMetIle 860
Db       2607 CTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATC 2666

Qy        861 LeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAlaAla 880
Db       2667 CTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCA 2726

Qy        881 ValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
Db       2727 GTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2783

RESULT 2
US-09-918-779-1
; Sequence 1, Application US/09918779
; Publication No. US20030064369A1
; GENERAL INFORMATION:
;  APPLICANT: Taupier, Raymond
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Spaderna, Steven
;  APPLICANT:  Shimkets, Richard
;  APPLICANT:  Zerhusen, Bryan
;  APPLICANT:  Spytek, Kimberly
;  APPLICANT:  Shenoy, Suresh
;  APPLICANT:  Li, Li
;  APPLICANT:  Gusev, Vladimir
;  APPLICANT:  Grosse, William
;  APPLICANT:  Alsobrook, John
;  APPLICANT:  Lepley, Denise
;  APPLICANT:  Burgess, Catherine
;  APPLICANT:  Gerlach, Valerie
;  APPLICANT:  Ellerman, Karen
;  APPLICANT:  MacDougall, John
;  APPLICANT:  Stone, David
;  APPLICANT:  Smithson, Glennda
;  TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
;  FILE REFERENCE: 21402-074 US
;  CURRENT APPLICATION NUMBER: US/09/918,779
;  CURRENT FILING DATE:  2001-07-30
;  PRIOR APPLICATION NUMBER: 60/221,409
;  PRIOR FILING DATE: 2000-07-28
;  PRIOR APPLICATION NUMBER: 60/222,840
;  PRIOR FILING DATE: 2000-08-04
;  PRIOR APPLICATION NUMBER: 60/223,752
;  PRIOR FILING DATE: 2000-08-08
;  PRIOR APPLICATION NUMBER: 60/223,762
;  PRIOR FILING DATE: 2000-08-08
;  PRIOR APPLICATION NUMBER: 60/223,770
;  PRIOR FILING DATE: 2000-08-08
;  PRIOR APPLICATION NUMBER: 60/223,769
;  PRIOR FILING DATE: 2000-08-08
;  PRIOR APPLICATION NUMBER: 60/225,146
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/225,392
;  PRIOR FILING DATE: 2000-08-15
;  PRIOR APPLICATION NUMBER: 60/225,470
;  PRIOR FILING DATE: 2000-08-15
;  PRIOR APPLICATION NUMBER: 60/225,697
;  PRIOR FILING DATE: 2000-08-16
;  PRIOR APPLICATION NUMBER: 60/263,662
;  PRIOR FILING DATE: 2001-02-01
;  PRIOR APPLICATION NUMBER: 60/281,645
;  PRIOR FILING DATE: 2001-04-05
;  NUMBER OF SEQ ID NOS: 61
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 2752
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-918-779-1

Alignment Scores:
Pred. No.:               0           Length:         2752
Score:                   4698.50     Matches:        888
Percent Similarity:      98.89%      Conservative:   2
Best Local Similarity:   98.67%      Mismatches:     7
Query Match:             98.15%      Indels:         3
DB:                      13          Gaps:           3

US-09-970-944-2 (1-899) x US-09-918-779-1 (1-2752)

Qy          1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeu 20
Db         46 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 105

Qy         21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40
Db        106 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 165

Qy         41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro 60
Db        166 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 225

Qy         61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80
Db        226 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 285

Qy         81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100
Db        286 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 345

Qy        101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120
Db        346 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 405

Qy        121 GluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAla 140
Db        406 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 465

Qy        141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160
Db        466 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 525
```

```
Qy    161 SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    526 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 585

Qy    181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    586 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 645

Qy    201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    646 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 705

Qy    221 CysValAlaLysAsnIleValAlaArgArgArgSerAlaSerAlaAlaValIleValTyr 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    706 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 765

Qy    241 ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg 260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    766 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 825

Qy    261 GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe 280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    826 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 885

Qy    281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300
          |||||||||||||||||          :::|||        :::     |||    |||||||||
Db    886 TGTGAGGGGCAGAATGTC---CAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGC 942

Qy    301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    943 AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC 1002

Qy    321 ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeu 340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1003 CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG 1062

Qy    341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1063 GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG 1122

Qy    361 AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuLeuValLeu 380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1123 GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC 1182

Qy    381 IleLeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIle 400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1183 ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1242

Qy    401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1243 CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1302

Qy    421 LeuThrIleGlnProAspLeuSer---ThrThrThrThrTyrGlnGlySerLeuCysPro 439
          ||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
Db   1303 CTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCC 1362

Qy    440 ArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu 459
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1363 CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG 1422

Qy    460 GlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheVal 479
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1423 GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC 1482

Qy    480 SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr 499
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1483 TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC 1542

Qy    500 TyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeu 519
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1543 TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC 1602

Qy    520 LeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHis 539
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1603 CTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC 1662

Qy    540 LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal 559
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1663 AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT 1722

Qy    560 SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys 579
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1723 AGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT 1782

Qy    580 GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySer 599
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1783 GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC 1842

Qy    600 TrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGln 619
          ||||||   |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1843 TGGGAG---GATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG 1899

Qy    620 LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly 639
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1900 CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA 1959

Qy    640 GluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAla 659
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1960 GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC 2019

Qy    660 CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu 679
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2020 TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC 2079

Qy    680 LysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgVal 699
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2080 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTC 2139

Qy    700 LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2140 CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCC 2199

Qy    720 LeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn 739
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2200 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2259

Qy    740 GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2260 GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2319

Qy    760 AspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIle 779
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2320 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2379

Qy    780 AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAla 799
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2380 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2439

Qy    800 GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2440 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2499

Qy    820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2500 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2559

Qy    840 LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2560 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG 2619

Qy    860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2620 ATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2679

Qy    880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       2680 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2739

RESULT 3
US-10-624-932-1
; Sequence 1, Application US/10624932
; Publication No. US20040096877A1
; GENERAL INFORMATION:
;  APPLICANT: Taupier, Raymond
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Spaderna, Steven
;  APPLICANT:  Shimkets, Richard
;  APPLICANT:  Zerhusen, Bryan
;  APPLICANT:  Spytek, Kimberly
;  APPLICANT:  Shenoy, Suresh
;  APPLICANT:  Li, Li
;  APPLICANT:  Gusev, Vladimir
;  APPLICANT:  Grosse, William
;  APPLICANT:  Alsobrook, John
;  APPLICANT:  Lepley, Denise
;  APPLICANT:  Burgess, Catherine
;  APPLICANT:  Gerlach, Valerie
;  APPLICANT:  Ellerman, Karen
;  APPLICANT:  MacDougall, John
;  APPLICANT:  Stone, David
;  APPLICANT:  Smithson, Glennda
;  TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
;  FILE REFERENCE: 21402-074 US
;  CURRENT APPLICATION NUMBER: US/10/624,932
;  CURRENT FILING DATE:  2003-07-21
;  PRIOR APPLICATION NUMBER: 09/918,779
;  PRIOR FILING DATE: 2001-07-03
;  PRIOR APPLICATION NUMBER: 60/221,409
;  PRIOR FILING DATE: 2000-07-28
;  PRIOR APPLICATION NUMBER: 60/222,840
;  PRIOR FILING DATE: 2000-08-04
;  PRIOR APPLICATION NUMBER: 60/223,752
;  PRIOR FILING DATE: 2000-08-08
;  PRIOR APPLICATION NUMBER: 60/223,762
;  PRIOR FILING DATE: 2000-08-08
;  PRIOR APPLICATION NUMBER: 60/223,770
;  PRIOR FILING DATE: 2000-08-08
;  PRIOR APPLICATION NUMBER: 60/223,769
;  PRIOR FILING DATE: 2000-08-08
;  PRIOR APPLICATION NUMBER: 60/225,146
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/225,392
;  PRIOR FILING DATE: 2000-08-15
;  PRIOR APPLICATION NUMBER: 60/225,470
;  PRIOR FILING DATE: 2000-08-15
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 61
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 2752
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-624-932-1

Alignment Scores:
Pred. No.:               0           Length:          2752
Score:                   4698.50     Matches:         888
Percent Similarity:      98.89%      Conservative:    2
Best Local Similarity:   98.67%      Mismatches:      7
Query Match:             98.15%      Indels:          3
DB:                      17          Gaps:            3

US-09-970-944-2 (1-899) x US-10-624-932-1 (1-2752)

Qy          1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         46 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 105

Qy         21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        106 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 165

Qy         41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        166 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 225

Qy         61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        226 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 285

Qy         81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        286 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 345

Qy        101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        346 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 405

Qy        121 GluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        406 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 465

Qy        141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        466 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 525

Qy        161 SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        526 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 585

Qy        181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        586 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 645

Qy        201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        646 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 705

Qy        221 CysValAlaLysAsnIleValAlaArgArgArgSerAlaSerAlaAlaValIleValTyr 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        706 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 765

Qy        241 ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        766 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 825

Qy        261 GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        826 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 885

Qy        281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300
              ||||||||||||||||||      :::|||      :::    |||      ||||||||
Db        886 TGTGAGGGGCAGAATGTC---CAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGC 942

Qy        301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        943 AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC 1002

Qy        321 ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1003 CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG 1062

Qy        341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1063 GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG 1122

Qy        361 AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuLeuValLeu 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1123 GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC 1182

Qy        381 IleLeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIle 400
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1183 ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1242

Qy        401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1243 CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1302

Qy        421 LeuThrIleGlnProAspLeuSer---ThrThrThrThrTyrGlnGlySerLeuCysPro 439
              ||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
Db       1303 CTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCC 1362

Qy        440 ArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu 459
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1363 CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG 1422

Qy        460 GlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheVal 479
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1423 GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC 1482

Qy        480 SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr 499
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1483 TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC 1542

Qy        500 TyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeu 519
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1543 TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC 1602

Qy        520 LeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHis 539
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1603 CTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC 1662

Qy        540 LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal 559
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1663 AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT 1722

Qy        560 SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys 579
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1723 AGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT 1782

Qy        580 GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySer 599
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1783 GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC 1842

Qy        600 TrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGln 619
              ||||||   |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1843 TGGGAG---GATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG 1899

Qy        620 LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly 639
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1900 CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA 1959

Qy        640 GluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAla 659
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1960 GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC 2019

Qy        660 CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu 679
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2020 TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC 2079

Qy        680 LysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgVal 699
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2080 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTC 2139

Qy        700 LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2140 CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCC 2199

Qy        720 LeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn 739
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2200 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2259

Qy        740 GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2260 GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2319

Qy        760 AspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIle 779
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2320 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2379

Qy        780 AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAla 799
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2380 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2439

Qy        800 GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2440 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2499

Qy        820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2500 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2559

Qy        840 LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2560 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG 2619

Qy        860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2620 ATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2679

Qy        880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2680 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2739
```

```
RESULT 4
US-10-240-154-15
; Sequence 15, Application US/10240154
; Publication No. US20030175741A1
; GENERAL INFORMATION:
;  APPLICANT: Cochran et al.
;  TITLE OF INVENTION: SCHIZOPHRENIA RELATED GENES
;  FILE REFERENCE: CKFW-P01-006
;  CURRENT APPLICATION NUMBER: US/10/240,154
;  CURRENT FILING DATE:  2001-04-02
;  PRIOR APPLICATION NUMBER: PCT/GB01/01486
;  PRIOR FILING DATE: 2001-04-02
;  NUMBER OF SEQ ID NOS: 34
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 15
;   LENGTH: 2697
;   TYPE: DNA
;   ORGANISM: Rattus sp.
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(2697)
US-10-240-154-15

Alignment Scores:
Pred. No.:              0           Length:         2697
Score:                  4545.50     Matches:        852
Percent Similarity:     96.78%      Conservative:   19
Best Local Similarity:  94.67%      Mismatches:     26
Query Match:            94.96%      Indels:         3
DB:                     15          Gaps:           3

US-09-970-944-2 (1-899) x US-10-240-154-15 (1-2697)

Qy          1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeu 20
              ||||||||||||||||||||||||||   |||||||||||||||||||||||||||||||
Db          1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGTGCTCCTGGGCATAGTCCTCGCCGCCTGGCTT 60

Qy         21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGTGGTTCGGGTGCCCAGCAGAGTGCCACGGTGGCCAATCCAGTGCCCGGTGCCAACCCC 120

Qy         41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    121 GACCTGCTGCCCCACTTCCTGGTAGAGCCTGAGGACGTGTACATTGTCAAGAACAAGCCG 180

Qy     61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GTGTTGTTGGTGTGCAAGGCTGTGCCTGCCACCCAGATCTTCTTCAAGTGCAATGGGGAA 240

Qy     81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100
          ||||||||||||||||||||||||||||||||||||||||||   |||||||||   |||
Db    241 TGGGTCCGCCAGGTCGATCACGTAATTGAACGCAGCACCGACAGCAGCAGCGGATTGCCA 300

Qy    101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 ACCATGGAGGTCCGTATCAACGTATCGAGGCAGCAGGTAGAGAAAGTGTTTGGGCTGGAG 360

Qy    121 GluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAla 140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GAATACTGGTGCCAGTGTGTGGCATGGAGCTCCTCGGGTACCACCAAAAGTCAGAAGGCC 420

Qy    141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160
          |||||||||||||||   ||||||||||||||||||||||||||||||||||||||||||
Db    421 TACATCCGGATTGCCTATTTGCGCAAGAACTTTGAGCAGGAGCCACTGGCCAAGGAAGTG 480

Qy    161 SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TCACTGGAGCAAGGCATTGTACTACCTTGTCGCCCCCCAGAAGGAATCCCCCCAGCTGAG 540

Qy    181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GTGGAGTGGCTTCGAAATGAGGACCTCGTGGACCCCTCCCTCGATCCCAATGTGTACATC 600

Qy    201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 ACGCGGGAGCACAGCCTAGTCGTGCGTCAGGCCCGCCTGGCCGACACGGCCAACTACACC 660

Qy    221 CysValAlaLysAsnIleValAlaArgArgArgSerAlaSerAlaAlaValIleValTyr 240
          ||||||||||||||||||||||||||||||||||||   |||||||||||||||||||||
Db    661 TGTGTGGCCAAGAACATCGTAGCCCGTCGCCGAAGCACCTCTGCAGCGGTCATTGTTTAT 720

Qy    241 ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg 260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GTGAACGGTGGGTGGTCGACGTGGACTGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGT 780

Qy    261 GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe 280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCACCTCTCAACGGGGGCGCCTTC 840

Qy    281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300
          ||||||||||||||||||      :::|||      :::   |||      |||||||||
Db    841 TGTGAGGGGCAGAATGTC---CAGAAAACAGCCTGCGCCACTCTGTGCCCAGTGGATGGG 897

Qy    301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320
          |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db    898 AGCTGGAGTTCGTGGAGTAAGTGGTCAGCCTGTGGGCTTGACTGCACCCACTGGCGGAGC 957

Qy    321 ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeu 340
          |||||||||||||||||||||||||||||||||||||||||||||:::|||   ||||||
Db    958 CGCGAGTGCTCTGACCCAGCACCCCGCAATGGAGGTGAGGAGTGTCGGGGTGCTGACCTG 1017

Qy    341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360
          ||||||||||||||||||||||||||||||:::|||:::||||||   ||||||||||||
Db   1018 GACACCCGCAACTGTACCAGTGACCTCTGCCTGCACACCGCTTCTTGCCCCGAGGACGTG 1077

Qy    361 AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuLeuValLeu 380
          |||||||||:::||||||:::||||||||||||||||||   ||||||||||||   |||
Db   1078 GCTCTCTACATCGGCCTTGTCGCTGTGGCTGTGTGCCTCTTCTTGCTGTTGCTGGCCCTT 1137

Qy    381 IleLeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIle 400
             |||:::|||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1138 GGACTCATTTACTGTCGCAAGAAGGAAGGGCTGGACTCCGATGTGGCCGACTCGTCCATC 1197

Qy    401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1198 CTCACCTCGGGCTTCCAGCCTGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCACCTG 1257

Qy    421 LeuThrIleGlnProAspLeuSer---ThrThrThrThrTyrGlnGlySerLeuCysPro 439
          ||||||||||||||||||||||||   ||||||||||||||||||||||||||||||   
Db   1258 CTCACCATCCAGCCAGACCTCAGCACCACCACTACCACCTACCAGGGCAGTCTATGTTCG 1317

Qy    440 ArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu 459
          |||||||||||||||||||||||||||||||||:::||||||||||||||||||||||||
Db   1318 AGGCAGGATGGACCCAGCCCCAAGTTCCAGCTCTCTAATGGTCACCTGCTCAGCCCACTG 1377

Qy    460 GlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheVal 479
          |||   |||||||||||||||||||||||||||||||||||||||||||||:::||||||
Db   1378 GGGAGTGGCCGCCATACGTTGCACCACAGCTCACCCACCTCTGAGGCTGAGGACTTCGTC 1437

Qy    480 SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr 499
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||   
Db   1438 TCCCGCCTCTCCACCCAAAACTACTTTCGTTCCCTGCCCCGCGGCACCAGCAACATGGCC 1497

Qy    500 TyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeu 519
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1498 TACGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACGGGGATCAGCCTC 1557

Qy    520 LeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHis 539
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1558 CTCATACCCCCGGATGCCATCCCCCGAGGAAAGATCTACGAGATCTACCTCACACTGCAC 1617

Qy    540 LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal 559
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db   1618 AAGCCAGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCAGTCGTT 1677

Qy    560 SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys 579
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1678 AGCTGTGGGCCCCCAGGAGTCCTGCTCACCCGGCCAGTCATCCTTGCAATGGACCACTGT 1737

Qy    580 GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySer 599
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1738 GGAGAGCCCAGCCCTGACAGCTGGAGTCTGCGCCTCAAAAAGCAGTCCTGCGAGGGCAGT 1797

Qy    600 TrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGln 619
          ||||||   |||||||||||||||||||||||||||:::|||||||||||||||||||||
Db   1798 TGGGAG---GATGTGCTGCACCTTGGTGAGGAGTCACCTTCCCACCTCTACTACTGCCAG 1854

Qy    620 LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly 639
          |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1855 CTGGAGGCCGGGGCCTGCTATGTCTTCACGGAGCAGCTGGGCCGCTTTGCCCTGGTAGGA 1914

Qy    640 GluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAla 659
          |||||||||||||||||||||   |||||||||:::||||||||||||||||||||||||
Db   1915 GAGGCCCTCAGCGTGGCTGCCACCAAGCGCCTCAGGCTCCTTCTGTTTGCTCCCGTGGCC 1974

Qy    660 CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu 679
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1975 TGTACGTCCCTTGAGTACAACATCCGAGTGTACTGCCTACACGACACCCACGACGCTCTC 2034

Qy    680 LysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgVal 699
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2035 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTAGGTGGACAGCTGATCCAGGAGCCTCGCGTC 2094

Qy    700 LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2095 CTGCACTTCAAAGACAGTTACCACAACCTACGTCTCTCCATCCACGACGTGCCCAGCTCC 2154

Qy    720 LeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn 739
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2155 CTGTGGAAGAGCAAGCTACTTGTCAGCTACCAGGAGATCCCTTTTTACCACATCTGGAAC 2214

Qy    740 GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759
          |||||||||:::||||||||||||||||||||||||||||||||::::::   |||||||||
Db   2215 GGCACCCAGCAGTATCTGCACTGCACCTTCACCCTGGAGCGCATCAACGCCAGCACCAGC 2274

Qy    760 AspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIle 779
          |||||||||||||||:::||||||||||||||||||||||||||||||||||||:::|||
Db   2275 GACCTGGCCTGCAAGGTGTGGGTGTGGCAGGTGGAGGGAGATGGGCAGAGCTTCAACATC 2334
```

```
Qy       780 AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAla 799
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2335 AACTTCAACATCACTAAGGACACAAGGTTTGCTGAATTGTTGGCTCTGGAGAGTGAAGGG 2394

Qy       800 GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2395 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAAAAG 2454

Qy       820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839
             ||||||:::||||||||||||||||||||||||   |||||||||||||||||||||||||
Db      2455 ATCATCGCCAGTCTGGACCCACCCTGCAGCCGGGGCGCCGACTGGAGAACTCTAGCCCAG 2514

Qy       840 LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2515 AAACTTCACCTGGACAGCCATCTTAGCTTCTTTGCCTCCAAGCCCAGCCCTACAGCCATG 2574

Qy       860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879
             ||||||||||||||||||||||||||||||||||||||||||   ||||||||||||
Db      2575 ATCCTCAACCTATGGGAGGCACGGCACTTCCCCAACGGCAACCTCGGCCAGCTGGCAGCA 2634

Qy       880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2635 GCTGTGGCCGGACTGGGCCAACCAGATGCTGGCCTCTTCACGGTGTCGGAGGCCGAGTGT 2694
```

RESULT 5
US-09-933-261-1
; Sequence 1, Application US/09933261
; Publication No. US20030040046A1
;   GENERAL INFORMATION:
;        APPLICANT: Tessier-Lavigne, Marc
;                   Leonardo, E. David
;                   Hink, Lindsay
;                   Masu, Masayuki
;                   Kazuko, Keino-Masu
;        TITLE OF INVENTION: Netrin Receptors
;        NUMBER OF SEQUENCES: 8
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
;             STREET: 268 BUSH STREET, SUITE 3200
;             CITY: SAN FRANCISCO
;             STATE: CALIFORNIA
;             COUNTRY: USA
;             ZIP: 94104
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/933,261
;             FILING DATE: 20-Aug-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/808,982
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: OSMAN, RICHARD A
;             REGISTRATION NUMBER: 36,627
;             REFERENCE/DOCKET NUMBER: UC96-217
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 343-4341
;             TELEFAX: (415) 343-4342
;   INFORMATION FOR SEQ ID NO: 1:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 3014 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-933-261-1

Alignment Scores:

```
Pred. No.:                 0            Length:          3014
Score:                     4545.50      Matches:         852
Percent Similarity:        96.78%       Conservative:    19
Best Local Similarity:     94.67%       Mismatches:      26
Query Match:               94.96%       Indels:          3
DB:                        10           Gaps:            3

US-09-970-944-2 (1-899) x US-09-933-261-1 (1-3014)

Qy         1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeu 20
             ||||||||||||||||||||||||   ||||||||||||||||||||||||||||||||
Db         1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGTGCTCCTGGGCATAGTCCTCGCCGCCTGGCTT 60

Qy        21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 CGTGGTTCGGGTGCCCAGCAGAGTGCCACGGTGGCCAATCCAGTGCCCGGTGCCAACCCC 120

Qy        41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 GACCTGCTGCCCCACTTCCTGGTAGAGCCTGAGGACGTGTACATTGTCAAGAACAAGCCG 180

Qy        61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 GTGTTGTTGGTGTGCAAGGCTGTGCCTGCCACCCAGATCTTCTTCAAGTGCAATGGGGAA 240

Qy        81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100
             ||||||||||||||||||||||||||||||||||||||||||   ||||||||   |||
Db       241 TGGGTCCGCCAGGTCGATCACGTAATTGAACGCAGCACCGACAGCAGCAGCGGATTGCCA 300

Qy       101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 ACCATGGAGGTCCGTATCAACGTATCGAGGCAGCAGGTAGAGAAAGTGTTTGGGCTGGAG 360

Qy       121 GluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAla 140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 GAATACTGGTGCCAGTGTGTGGCATGGAGCTCCTCGGGTACCACCAAAAGTCAGAAGGCC 420

Qy       141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160
             |||||||||||||||    |||||||||||||||||||||||||||||||||||||||||
Db       421 TACATCCGGATTGCCTATTTGCGCAAGAACTTTGAGCAGGAGCCACTGGCCAAGGAAGTG 480

Qy       161 SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 TCACTGGAGCAAGGCATTGTACTACCTTGTCGCCCCCCAGAAGGAATCCCCCCAGCTGAG 540

Qy       181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       541 GTGGAGTGGCTTCGAAATGAGGACCTCGTGGACCCCTCCCTCGATCCCAATGTGTACATC 600

Qy       201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       601 ACGCGGGAGCACAGCCTAGTCGTGCGTCAGGCCCGCCTGGCCGACACGGCCAACTACACC 660

Qy       221 CysValAlaLysAsnIleValAlaArgArgArgSerAlaSerAlaAlaValIleValTyr 240
             ||||||||||||||||||||||||||||||||||||   ||||||||||||||||||||
Db       661 TGTGTGGCCAAGAACATCGTAGCCCGTCGCCGAAGCACCTCTGCAGCGGTCATTGTTTAT 720

Qy       241 ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg 260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       721 GTGAACGGTGGGTGGTCGACGTGGACTGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGT 780

Qy       261 GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe 280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       781 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCACCTCTCAACGGGGGCGCCTTC 840

Qy       281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300
             |||||||||||||||||      :::|||       :::    |||      ||||||||
Db       841 TGTGAGGGGCAGAATGTC---CAGAAAACAGCCTGCGCCACTCTGTGCCCAGTGGATGGG 897

Qy       301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320
             ||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db       898 AGCTGGAGTTCGTGGAGTAAGTGGTCAGCCTGTGGGCTTGACTGCACCCACTGGCGGAGC 957
```

```
Qy      321 ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeu 340
Db      958 CGCGAGTGCTCTGACCCAGCACCCCGCAATGGAGGTGAGGAGTGTCGGGGTGCTGACCTG 1017

Qy      341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360
Db     1018 GACACCCGCAACTGTACCAGTGACCTCTGCCTGCACACCGCTTCTTGCCCCGAGGACGTG 1077

Qy      361 AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuValLeu 380
Db     1078 GCTCTCTACATCGGCCTTGTCGCTGTGGCTGTGTGCCTCTTCTTGCTGTTGCTGGCCCTT 1137

Qy      381 IleLeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIle 400
Db     1138 GGACTCATTTACTGTCGCAAGAAGGAAGGGCTGGACTCCGATGTGGCCGACTCGTCCATC 1197

Qy      401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420
Db     1198 CTCACCTCGGGCTTCCAGCCTGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCACCTG 1257

Qy      421 LeuThrIleGlnProAspLeuSer---ThrThrThrThrTyrGlnGlySerLeuCysPro 439
Db     1258 CTCACCATCCAGCCAGACCTCAGCACCACCACTACCACCTACCAGGGCAGTCTATGTTCG 1317

Qy      440 ArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu 459
Db     1318 AGGCAGGATGGACCCAGCCCCAAGTTCCAGCTCTCTAATGGTCACCTGCTCAGCCCACTG 1377

Qy      460 GlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheVal 479
Db     1378 GGGAGTGGCCGCCATACGTTGCACCACAGCTCACCCACCTCTGAGGCTGAGGACTTCGTC 1437

Qy      480 SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr 499
Db     1438 TCCCGCCTCTCCACCCAAAACTACTTTCGTTCCCTGCCCCGCGGCACCAGCAACATGGCC 1497

Qy      500 TyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeu 519
Db     1498 TACGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACGGGGATCAGCCTC 1557

Qy      520 LeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHis 539
Db     1558 CTCATACCCCCGGATGCCATCCCCCGAGGAAAGATCTACGAGATCTACCTCACACTGCAC 1617

Qy      540 LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal 559
Db     1618 AAGCCAGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCAGTCGTT 1677

Qy      560 SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys 579
Db     1678 AGCTGTGGGCCCCCAGGAGTCCTGCTCACCCGGCCAGTCATCCTTGCAATGGACCACTGT 1737

Qy      580 GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySer 599
Db     1738 GGAGAGCCCAGCCCTGACAGCTGGAGTCTGCGCCTCAAAAAGCAGTCCTGCGAGGGCAGT 1797

Qy      600 TrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGln 619
Db     1798 TGGGAG---GATGTGCTGCACCTTGGTGAGGAGTCACCTTCCCACCTCTACTACTGCCAG 1854

Qy      620 LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly 639
Db     1855 CTGGAGGCCGGGGCCTGCTATGTCTTCACGGAGCAGCTGGGCCGCTTTGCCCTGGTAGGA 1914

Qy      640 GluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAla 659
Db     1915 GAGGCCCTCAGCGTGGCTGCCACCAAGCGCCTCAGGCTCCTTCTGTTTGCTCCCGTGGCC 1974

Qy      660 CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu 679
Db     1975 TGTACGTCCCTTGAGTACAACATCCGAGTGTACTGCCTACACGACACCCACGACGCTCTC 2034

Qy      680 LysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgVal 699
Db     2035 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTAGGTGGACAGCTGATCCAGGAGCCTCGCGTC 2094

Qy      700 LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719
Db     2095 CTGCACTTCAAAGACAGTTACCACAACCTACGTCTCTCCATCCACGACGTGCCCAGCTCC 2154

Qy      720 LeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn 739
Db     2155 CTGTGGAAGAGCAAGCTACTTGTCAGCTACCAGGAGATCCCTTTTTACCACATCTGGAAC 2214

Qy      740 GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759
Db     2215 GGCACCCAGCAGTATCTGCACTGCACCTTCACCCTGGAGCGCATCAACGCCAGCACCAGC 2274

Qy      760 AspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIle 779
Db     2275 GACCTGGCCTGCAAGGTGTGGGTGTGGCAGGTGGAGGGAGATGGGCAGAGCTTCAACATC 2334

Qy      780 AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAla 799
Db     2335 AACTTCAACATCACTAAGGACACAAGGTTTGCTGAATTGTTGGCTCTGGAGAGTGAAGGG 2394

Qy      800 GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
Db     2395 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAAAAG 2454

Qy      820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839
Db     2455 ATCATCGCCAGTCTGGACCCACCCTGCAGCCGGGGCGCCGACTGGAGAACTCTAGCCCAG 2514

Qy      840 LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859
Db     2515 AAACTTCACCTGGACAGCCATCTTAGCTTCTTTGCCTCCAAGCCCAGCCCTACAGCCATG 2574

Qy      860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879
Db     2575 ATCCTCAACCTATGGGAGGCACGGCACTTCCCCAACGGCAACCTCGGCCAGCTGGCAGCA 2634

Qy      880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
Db     2635 GCTGTGGCCGGACTGGGCCAACCAGATGCTGGCCTCTTCACGGTGTCGGAGGCCGAGTGT 2694
```

RESULT 6
US-10-256-702-1
; Sequence 1, Application US/10256702
; Publication No. US20030059859A1
;   GENERAL INFORMATION:
;        APPLICANT: Tessier-Lavigne, Marc
;                   Leonardo, E. David
;                   Hink, Lindsay
;                   Masu, Masayuki
;                   Kazuko, Keino-Masu
;        TITLE OF INVENTION: Netrin Receptors
;        NUMBER OF SEQUENCES: 8
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
;             STREET: 268 BUSH STREET, SUITE 3200
;             CITY: SAN FRANCISCO
;             STATE: CALIFORNIA
;             COUNTRY: USA
;             ZIP: 94104
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/256,702
;             FILING DATE: 27-Sep-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/933,261

```
;            FILING DATE: 20-Aug-2001
;            APPLICATION NUMBER: 08/808,982
;            FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;            NAME: OSMAN, RICHARD A
;            REGISTRATION NUMBER: 36,627
;            REFERENCE/DOCKET NUMBER: UC96-217
;        TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (415) 343-4341
;            TELEFAX: (415) 343-4342
;    INFORMATION FOR SEQ ID NO: 1:
;        SEQUENCE CHARACTERISTICS:
;            LENGTH: 3014 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: double
;            TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-256-702-1

Alignment Scores:
Pred. No.:              0             Length:         3014
Score:                  4545.50       Matches:        852
Percent Similarity:     96.78%        Conservative:   19
Best Local Similarity:  94.67%        Mismatches:     26
Query Match:            94.96%        Indels:         3
DB:                     15            Gaps:           3

US-09-970-944-2 (1-899) x US-10-256-702-1 (1-3014)

Qy        1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeu 20
Db        1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGTGCTCCTGGGCATAGTCCTCGCCGCCTGGCTT 60

Qy       21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40
Db       61 CGTGGTTCGGGTGCCCAGCAGAGTGCCACGGTGGCCAATCCAGTGCCCGGTGCCAACCCC 120

Qy       41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro 60
Db      121 GACCTGCTGCCCCACTTCCTGGTAGAGCCTGAGGACGTGTACATTGTCAAGAACAAGCCG 180

Qy       61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80
Db      181 GTGTTGTTGGTGTGCAAGGCTGTGCCTGCCACCCAGATCTTCTTCAAGTGCAATGGGGAA 240

Qy       81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100
Db      241 TGGGTCCGCCAGGTCGATCACGTAATTGAACGCAGCACCGACAGCAGCAGCGGATTGCCA 300

Qy      101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120
Db      301 ACCATGGAGGTCCGTATCAACGTATCGAGGCAGCAGGTAGAGAAAGTGTTTGGGCTGGAG 360

Qy      121 GluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAla 140
Db      361 GAATACTGGTGCCAGTGTGTGGCATGGAGCTCCTCGGGTACCACCAAAAGTCAGAAGGCC 420

Qy      141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160
Db      421 TACATCCGGATTGCCTATTTGCGCAAGAACTTTGAGCAGGAGCCACTGGCCAAGGAAGTG 480

Qy      161 SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu 180
Db      481 TCACTGGAGCAAGGCATTGTACTACCTTGTCGCCCCCCAGAAGGAATCCCCCCAGCTGAG 540

Qy      181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200
Db      541 GTGGAGTGGCTTCGAAATGAGGACCTCGTGGACCCCTCCCTCGATCCCAATGTGTACATC 600

Qy      201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220
Db      601 ACGCGGGAGCACAGCCTAGTCGTGCGTCAGGCCCGCCTGGCCGACACGGCCAACTACACC 660

Qy      221 CysValAlaLysAsnIleValAlaArgArgArgSerAlaSerAlaAlaValIleValTyr 240
Db      661 TGTGTGGCCAAGAACATCGTAGCCCGTCGCCGAAGCACCTCTGCAGCGGTCATTGTTTAT 720

Qy      241 ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg 260
Db      721 GTGAACGGTGGGTGGTCGACGTGGACTGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGT 780

Qy      261 GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe 280
Db      781 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCACCTCTCAACGGGGGCGCCTTC 840

Qy      281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300
Db      841 TGTGAGGGGCAGAATGTC---CAGAAAACAGCCTGCGCCACTCTGTGCCCAGTGGATGGG 897

Qy      301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320
Db      898 AGCTGGAGTTCGTGGAGTAAGTGGTCAGCCTGTGGGCTTGACTGCACCCACTGGCGGAGC 957

Qy      321 ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeu 340
Db      958 CGCGAGTGCTCTGACCCAGCACCCCGCAATGGAGGTGAGGAGTGTCGGGGTGCTGACCTG 1017

Qy      341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360
Db     1018 GACACCCGCAACTGTACCAGTGACCTCTGCCTGCACACCGCTTCTTGCCCCGAGGACGTG 1077

Qy      361 AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuLeuValLeu 380
Db     1078 GCTCTCTACATCGGCCTTGTCGCTGTGGCTGTGTGCCTCTTCTTGCTGTTGCTGGCCCTT 1137

Qy      381 IleLeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIle 400
Db     1138 GGACTCATTTACTGTCGCAAGAAGGAAGGGCTGGACTCCGATGTGGCCGACTCGTCCATC 1197

Qy      401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420
Db     1198 CTCACCTCGGGCTTCCAGCCTGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCACCTG 1257

Qy      421 LeuThrIleGlnProAspLeuSer---ThrThrThrThrTyrGlnGlySerLeuCysPro 439
Db     1258 CTCACCATCCAGCCAGACCTCAGCACCACCACTACCACCTACCAGGGCAGTCTATGTTCG 1317

Qy      440 ArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu 459
Db     1318 AGGCAGGATGGACCCAGCCCCAAGTTCCAGCTCTCTAATGGTCACCTGCTCAGCCCACTG 1377

Qy      460 GlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheVal 479
Db     1378 GGGAGTGGCCGCCATACGTTGCACCACAGCTCACCCACCTCTGAGGCTGAGGACTTCGTC 1437

Qy      480 SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr 499
Db     1438 TCCCGCCTCTCCACCCAAAACTACTTTCGTTCCCTGCCCCGCGGCACCAGCAACATGGCC 1497

Qy      500 TyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeu 519
Db     1498 TACGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACGGGGATCAGCCTC 1557

Qy      520 LeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHis 539
Db     1558 CTCATACCCCCGGATGCCATCCCCCGAGGAAAGATCTACGAGATCTACCTCACACTGCAC 1617

Qy      540 LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal 559
Db     1618 AAGCCAGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCAGTCGTT 1677

Qy      560 SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys 579
Db     1678 AGCTGTGGGCCCCCAGGAGTCCTGCTCACCCGGCCAGTCATCCTTGCAATGGACCACTGT 1737
```

```
Qy    580 GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySer 599
Db   1738 GGAGAGCCCAGCCCTGACAGCTGGAGTCTGCGCCTCAAAAAGCAGTCCTGCGAGGGCAGT 1797

Qy    600 TrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGln 619
Db   1798 TGGGAG---GATGTGCTGCACCTTGGTGAGGAGTCACCTTCCCACCTCTACTACTGCCAG 1854

Qy    620 LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly 639
Db   1855 CTGGAGGCCGGGGCCTGCTATGTCTTCACGGAGCAGCTGGGCCGCTTTGCCCTGGTAGGA 1914

Qy    640 GluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAla 659
Db   1915 GAGGCCCTCAGCGTGGCTGCCACCAAGCGCCTCAGGCTCCTTCTGTTTGCTCCCGTGGCC 1974

Qy    660 CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu 679
Db   1975 TGTACGTCCCTTGAGTACAACATCCGAGTGTACTGCCTACACGACACCCACGACGCTCTC 2034

Qy    680 LysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgVal 699
Db   2035 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTAGGTGGACAGCTGATCCAGGAGCCTCGCGTC 2094

Qy    700 LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719
Db   2095 CTGCACTTCAAAGACAGTTACCACAACCTACGTCTCTCCATCCACGACGTGCCCAGCTCC 2154

Qy    720 LeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn 739
Db   2155 CTGTGGAAGAGCAAGCTACTTGTCAGCTACCAGGAGATCCCTTTTTACCACATCTGGAAC 2214

Qy    740 GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759
Db   2215 GGCACCCAGCAGTATCTGCACTGCACCTTCACCCTGGAGCGCATCAACGCCAGCACCAGC 2274

Qy    760 AspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIle 779
Db   2275 GACCTGGCCTGCAAGGTGTGGGTGTGGCAGGTGGAGGGAGATGGGCAGAGCTTCAACATC 2334

Qy    780 AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAla 799
Db   2335 AACTTCAACATCACTAAGGACACAAGGTTTGCTGAATTGTTGGCTCTGGAGAGTGAAGGG 2394

Qy    800 GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
Db   2395 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAAAAG 2454

Qy    820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839
Db   2455 ATCATCGCCAGTCTGGACCCACCCTGCAGCCGGGGCGCCGACTGGAGAACTCTAGCCCAG 2514

Qy    840 LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859
Db   2515 AAACTTCACCTGGACAGCCATCTTAGCTTCTTTGCCTCCAAGCCCAGCCCTACAGCCATG 2574

Qy    860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879
Db   2575 ATCCTCAACCTATGGGAGGCACGGCACTTCCCCAACGGCAACCTCGGCCAGCTGGCAGCA 2634

Qy    880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
Db   2635 GCTGTGGCCGGACTGGGCCAACCAGATGCTGGCCTCTTCACGGTGTCGGAGGCCGAGTGT 2694
```

RESULT 7
US-10-311-623-13
; Sequence 13, Application US/10311623
; Publication No. US20040023244A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; GRIFFIN, Jennifer A.
; APPLICANT: KALLICK, Deborah A.; TRIBOULEY, Catherine M.
; APPLICANT: YUE, Henry; NGUYEN, Danniel B.
; APPLICANT: TANG, Y. Tom; LAL, Preeti G.
; APPLICANT: POLICKY, Jennifer L.; AZIMZAI, Yalda
; APPLICANT: LU, Dyung Aina M.; GRAUL, Richard C.
; APPLICANT: YAO, Monique G.; BURFORD, Neil
; APPLICANT: HAFALIA, April J. A.; BAUGHN, Mariah R.
; APPLICANT: BANDMAN, Olga; ARVIZU, Chandra S.
; APPLICANT: YANG, Junming; XU, Yuming
; APPLICANT: GANDHI, Ameena R.; WARREN, Bridget A.
; APPLICANT: DING, Li; SANJANWALA, Madhusudan M.
; APPLICANT: DUGGAN, Brendan M.; LU, Yan
; TITLE OF INVENTION: RECEPTORS
; FILE REFERENCE: PF-0793 USN
; CURRENT APPLICATION NUMBER: US/10/311,623
; CURRENT FILING DATE: 2002-12-17
; PRIOR APPLICATION NUMBER: US 01/19942
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/214,027
; PRIOR FILING DATE: 2000-06-21
; PRIOR APPLICATION NUMBER: US 60/228,045
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: US 60/255,104
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PERL Program
; SEQ ID NO 13
; LENGTH: 3580
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20040023244A1 6052371CB1
US-10-311-623-13

Alignment Scores:

| | | | |
|---|---|---|---|
| Pred. No.: | 0 | Length: | 3580 |
| Score: | 4379.50 | Matches: | 838 |
| Percent Similarity: | 93.22% | Conservative: | 1 |
| Best Local Similarity: | 93.11% | Mismatches: | 2 |
| Query Match: | 91.49% | Indels: | 59 |
| DB: | 17 | Gaps: | 3 |

US-09-970-944-2 (1-899) x US-10-311-623-13 (1-3580)

```
Qy      1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeu 20
Db      4 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 63

Qy     21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40
Db     64 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 123

Qy     41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro 60
Db    124 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 183

Qy     61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80
Db    184 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 243

Qy     81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100
Db    244 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 303

Qy    101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120
Db    304 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 363

Qy    121 GluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAla 140
Db    364 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 423

Qy    141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160
Db    424 TACATCCGCATAGCCTATTTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 483
```

```
Qy     161 SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     484 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 543

Qy     181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     544 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 603

Qy     201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     604 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 663

Qy     221 CysValAlaLysAsnIleValAlaArgArgArgSerAlaSerAlaAlaValIleValTyr 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     664 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 723

Qy     241 ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg 260

Db     723 ------------------------------------------------------------ 723

Qy     261 GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe 280

Db     723 ------------------------------------------------------------ 723

Qy     281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300
                                                                  |||||||||
Db     724 ---------------------------------------------------GTGGACGGC 732

Qy     301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     733 AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC 792

Qy     321 ArgGluCysSerAspProAlaProArgAsnGlyGlyGluCysGlnGlyThrAspLeu 340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     793 CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG 852

Qy     341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360
           ||||||||||||||||||||||||||||||||||||:::|||||||||||||||||||||
Db     853 GACACCCGCAACTGTACCAGTGACCTCTGTGTACACACTGCTTCTGGCCCTGAGGACGTG 912

Qy     361 AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuLeuValLeu 380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     913 GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC 972

Qy     381 IleLeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIle 400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     973 ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1032

Qy     401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1033 CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1092

Qy     421 LeuThrIleGlnProAspLeuSer---ThrThrThrThrTyrGlnGlySerLeuCysPro 439
           ||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
Db    1093 CTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCC 1152

Qy     440 ArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu 459
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1153 CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG 1212

Qy     460 GlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheVal 479
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1213 GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC 1272

Qy     480 SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr 499
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1273 TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC 1332

Qy     500 TyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeu 519
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1333 TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGAATCAGCCTC 1392

Qy     520 LeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHis 539
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1393 CTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC 1452

Qy     540 LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal 559
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1453 AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT 1512

Qy     560 SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys 579
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1513 AGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT 1572

Qy     580 GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySer 599
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1573 GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC 1632

Qy     600 TrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGln 619
           ||||||   |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1633 TGGGAG---GATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG 1689

Qy     620 LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly 639
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1690 CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA 1749

Qy     640 GluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAla 659
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1750 GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC 1809

Qy     660 CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu 679
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1810 TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC 1869

Qy     680 LysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgVal 699
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1870 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTC 1929

Qy     700 LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1930 CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCC 1989

Qy     720 LeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn 739
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1990 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2049

Qy     740 GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2050 GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2109

Qy     760 AspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIle 779
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2110 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2169

Qy     780 AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAla 799
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2170 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2229

Qy     800 GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2230 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2289

Qy     820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2290 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2349

Qy     840 LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2350 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG 2409

Qy     860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2410 ATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2469

Qy     880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
```

```
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2470 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2529

RESULT 8
US-09-933-261-2
; Sequence 2, Application US/09933261
; Publication No. US20030040046A1
;  GENERAL INFORMATION:
;       APPLICANT: Tessier-Lavigne, Marc
;                  Leonardo, E. David
;                  Hink, Lindsay
;                  Masu, Masayuki
;                  Kazuko, Keino-Masu
;       TITLE OF INVENTION: Netrin Receptors
;       NUMBER OF SEQUENCES: 8
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
;            STREET: 268 BUSH STREET, SUITE 3200
;            CITY: SAN FRANCISCO
;            STATE: CALIFORNIA
;            COUNTRY: USA
;            ZIP: 94104
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/933,261
;            FILING DATE: 20-Aug-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 08/808,982
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: OSMAN, RICHARD A
;            REGISTRATION NUMBER: 36,627
;            REFERENCE/DOCKET NUMBER: UC96-217
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (415) 343-4341
;            TELEFAX: (415) 343-4342
;  INFORMATION FOR SEQ ID NO: 2:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 1787 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: double
;            TOPOLOGY: linear
;       MOLECULE TYPE: cDNA
;       SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-933-261-2

Alignment Scores:
Pred. No.:              5.35e-274     Length:         1787
Score:                  2694.00       Matches:        543
Percent Similarity:     97.14%        Conservative:   1
Best Local Similarity:  96.96%        Mismatches:     11
Query Match:            56.28%        Indels:         11
DB:                     10            Gaps:           2

US-09-970-944-2 (1-899) x US-09-933-261-2 (1-1787)

Qy       344 AsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspValAlaLeuTyr 363
             ||||||||||||||||||   ||||||:::||||||||||||||||||||||||||||||
Db         3 AACTGTACCAGTGACCTCTG-GTACACACTGCTTCTGGCCCTGAGGACGTGGCCCTCTAT 61

Qy       364 ValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuLeuValLeuIleLeuVal 383
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        62 GTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCATCCTCGTT 121

Qy       384 TyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIleLeuThrSer 403
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       122 TATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCA 181

Qy       404 GlyPheGlnProValSerIle-LysProSerLysAlaAspAsnProHisLeuLeuThrIl 423
             |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db       182 GGCTTCCAGCCCGTCAGCATCTAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCAT 241

Qy       423 eGlnProAspLeuSer---ThrThrThrThrTyrGlnGlySerLeuCysProArgGlnAs 442
             ||||||||||||||||   |||||||||||||||||||||||||||||||||||||||||
Db       242 CCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGA 301

Qy       442 pGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeuGlyGlyGl 462
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       302 TGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGG 361

Qy       462 yArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheValSerArgLe 482
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       362 CCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCT 421

Qy       482 uSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThrTyrGlyTh 502
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       422 CTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGAC 481

Qy       502 rPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeuLeuIlePr 522
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       482 CTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGAATCAGCCTCCTCATCCC 541

Qy       522 oProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHisLysProGl 542
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       542 CCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGA 601

Qy       542 uAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleValSerCysGl 562
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       602 AGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGG 661

Qy       562 yProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGlyGluPr 582
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       662 ACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCC 721

Qy       582 oSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySerTrpGluGl 602
             ||||||||||||||||||||   |||||||||||||||||||||    ||||||
Db       722 CAGCCCTGACAGCTGGAGCCTGGCCCTCAAAAAGCAGTCGTGCGAGGGAGC-TGGGAG-- 778

Qy       602 nAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGlnLeuGluAl 622
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       779 -GATGT-CTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGC 836

Qy       622 aSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGluAlaLe 642
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       837 CAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCT 896

Qy       642 uSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAlaCysThrSe 662
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       897 CAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTC 956

Qy       662 rLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuLysGluVa 682
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       957 CCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGT 1016

Qy       682 lValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgValLeuHisPh 702
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1017 GGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTT 1076

Qy       702 eLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSerLeuTrpLy 722
              |||||||||||||||||||||   |||  ||||||||||||||||||||||||||||||
Db      1077 -AAGGACAGTTACCACAACCTGCC-CTATC-ATCCACGATGTGCCCAGCTCCCTGTGGAA 1133

Qy       722 sSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsnGlyThrGl 742
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1134 GAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCA 1193

Qy       742 nArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSerAspLeuAl 762
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1194 GCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGC 1253

Qy       762 aCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIleAsnPheAs 782
```

```
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1254 CTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAA 1313

Qy         782 nIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAlaGlyValPr 802
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1314 CATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCC 1373

Qy         802 oAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLysIleIleSe 822
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1374 AGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTC 1433

Qy         822 rSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGlnLysLeuHi 842
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1434 CAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCA 1493

Qy         842 sLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMetIleLeuAs 862
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1494 CCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAA 1553

Qy         862 nLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAlaAlaValAl 882
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1554 CCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGC 1613

Qy         882 aGlyLeuGlyGlnProAspAlaGlyLeu-PheThrVal-SerGluAlaGluCys 899
               ||||   |||||   |||    |||     |||||||| ||||||||||||||
Db        1614 TGGGACTGGCCAGCAGGACGGTGGCTTCTTTCACAGTGTTCGGAGGCTGAGTGC 1667

RESULT 9
US-10-256-702-2
; Sequence 2, Application US/10256702
; Publication No. US20030059859A1
;  GENERAL INFORMATION:
;       APPLICANT: Tessier-Lavigne, Marc
;                  Leonardo, E. David
;                  Hink, Lindsay
;                  Masu, Masayuki
;                  Kazuko, Keino-Masu
;       TITLE OF INVENTION: Netrin Receptors
;       NUMBER OF SEQUENCES: 8
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
;            STREET: 268 BUSH STREET, SUITE 3200
;            CITY: SAN FRANCISCO
;            STATE: CALIFORNIA
;            COUNTRY: USA
;            ZIP: 94104
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/256,702
;            FILING DATE: 27-Sep-2002
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/09/933,261
;            FILING DATE: 20-Aug-2001
;            APPLICATION NUMBER: 08/808,982
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: OSMAN, RICHARD A
;            REGISTRATION NUMBER: 36,627
;            REFERENCE/DOCKET NUMBER: UC96-217
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (415) 343-4341
;            TELEFAX: (415) 343-4342
;   INFORMATION FOR SEQ ID NO: 2:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 1787 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: double
;            TOPOLOGY: linear
;       MOLECULE TYPE: cDNA
;       SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-256-702-2

Alignment Scores:
Pred. No.:                5.35e-274    Length:        1787
Score:                    2694.00      Matches:       543
Percent Similarity:       97.14%       Conservative:  1
Best Local Similarity:    96.96%       Mismatches:    11
Query Match:              56.28%       Indels:        11
DB:                       15           Gaps:          2

US-09-970-944-2 (1-899) x US-10-256-702-2 (1-1787)

Qy         344 AsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspValAlaLeuTyr 363
               ||||||||||||||||||    |||||:::||||||||||||||||||||||||||||||
Db           3 AACTGTACCAGTGACCTCTG-GTACACACTGCTTCTGGCCCTGAGGACGTGGCCCTCTAT 61

Qy         364 ValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuLeuValLeuIleLeuVal 383
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          62 GTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCATCCTCGTT 121

Qy         384 TyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIleLeuThrSer 403
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         122 TATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCA 181

Qy         404 GlyPheGlnProValSerIle-LysProSerLysAlaAspAsnProHisLeuLeuThrIl 423
               ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db         182 GGCTTCCAGCCCGTCAGCATCTAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCAT 241

Qy         423 eGlnProAspLeuSer---ThrThrThrThrTyrGlnGlySerLeuCysProArgGlnAs 442
               ||||||||||||||||   |||||||||||||||||||||||||||||||||||||||||
Db         242 CCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGA 301

Qy         442 pGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeuGlyGlyGl 462
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         302 TGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGG 361

Qy         462 yArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheValSerArgLe 482
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         362 CCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCT 421

Qy         482 uSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThrTyrGlyTh 502
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         422 CTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGAC 481

Qy         502 rPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeuLeuIlePr 522
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         482 CTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGAATCAGCCTCCTCATCCC 541

Qy         522 oProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHisLysProGl 542
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         542 CCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGA 601

Qy         542 uAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleValSerCysGl 562
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         602 AGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGG 661

Qy         562 yProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGlyGluPr 582
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         662 ACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCC 721

Qy         582 oSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySerTrpGluGl 602
               |||||||||||||||||||||||||    ||||||||||||||||||||||  ||||||
Db         722 CAGCCCTGACAGCTGGAGCCTGGCCCTCAAAAAGCAGTCGTGCGAGGGAGC-TGGGAG-- 778

Qy         602 nAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGlnLeuGluAl 622
                |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         779 -GATGT-CTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGC 836

Qy         622 aSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGluAlaLe 642
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         837 CAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCT 896
```

```
Qy        642 uSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAlaCysThrSe 662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        897 CAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTC 956

Qy        662 rLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuLysGluVa 682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        957 CCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGT 1016

Qy        682 lValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgValLeuHisPh 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1017 GGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTT 1076

Qy        702 eLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSerLeuTrpLy 722
               ||||||||||||||||||||    |||    ||||||||||||||||||||||||||||
Db       1077 -AAGGACAGTTACCACAACCTGCC-CTATC-ATCCACGATGTGCCCAGCTCCCTGTGGAA 1133

Qy        722 sSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsnGlyThrGl 742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1134 GAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCA 1193

Qy        742 nArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSerAspLeuAl 762
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1194 GCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGC 1253

Qy        762 aCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIleAsnPheAs 782
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1254 CTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAA 1313

Qy        782 nIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAlaGlyValPr 802
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1314 CATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCC 1373

Qy        802 oAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLysIleIleSe 822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1374 AGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTC 1433

Qy        822 rSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGlnLysLeuHi 842
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1434 CAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCA 1493

Qy        842 sLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMetIleLeuAs 862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1494 CCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAA 1553

Qy        862 nLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAlaAlaValAl 882
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1554 CCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGC 1613

Qy        882 aGlyLeuGlyGlnProAspAlaGlyLeu-PheThrVal-SerGluAlaGluCys 899
              ||||    ||||||   |||    |||     ||||||||| ||||||||||||||
Db       1614 TGGGACTGGCCAGCAGGACGGTGGCTTCTTTCACAGTGTTCGGAGGCTGAGTGC 1667
```

RESULT 10
US-09-972-211-55
; Sequence 55, Application US/09972211
; Publication No. US20040048245A1
; GENERAL INFORMATION:
;  APPLICANT: Shimkets, Richard A
;  APPLICANT:  Taupier Jr, Raymond J
;  APPLICANT:  Burgess, Catherine E
;  APPLICANT:  Zerhusen, Bryan D
;  APPLICANT:  Mezes, Peter S
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Malyankar, Uriel M
;  APPLICANT:  Grosse, William M
;  APPLICANT:  Alsobrook II, John P
;  APPLICANT:  Lepley, Denise M
;  APPLICANT:  Spytek, Kimberly Ann
;  APPLICANT:  Li, Li
;  APPLICANT:  Edinger, Shlomit
;  APPLICANT:  Gerlach, Valerie
;  APPLICANT:  Ellerman, Karen
;  APPLICANT:  MacDougall, John R
;  APPLICANT:  Gunther, Erik
;  APPLICANT:  Millet, Isabelle
;  APPLICANT:  Stone, David J
;  APPLICANT:  Smithson, Glennda
;  APPLICANT:  Szekeres Jr, Edward S
;  TITLE OF INVENTION: No. US20040048245A1el Human Proteins, Polynucleotides Encoding Th
;  TITLE OF INVENTION:  Methods Of Using The Same
;  FILE REFERENCE: 21402-141
;  CURRENT APPLICATION NUMBER: US/09/972,211
;  CURRENT FILING DATE:  2001-10-05
;  PRIOR APPLICATION NUMBER: 60/238,325
;  PRIOR FILING DATE: 2000-10-05
;  PRIOR APPLICATION NUMBER: 60/238,323
;  PRIOR FILING DATE: 2000-10-05
;  PRIOR APPLICATION NUMBER: 60/238,400
;  PRIOR FILING DATE: 2000-10-06
;  PRIOR APPLICATION NUMBER: 60/238,397
;  PRIOR FILING DATE: 2000-10-06
;  PRIOR APPLICATION NUMBER: 60/238,401
;  PRIOR FILING DATE: 2000-10-06
;  PRIOR APPLICATION NUMBER: 60/238,379
;  PRIOR FILING DATE: 2000-10-06
;  PRIOR APPLICATION NUMBER: 60/238,402
;  PRIOR FILING DATE: 2000-10-06
;  PRIOR APPLICATION NUMBER: 30/238,384
;  PRIOR FILING DATE: 2000-10-06
;  PRIOR APPLICATION NUMBER: 60/238,373
;  PRIOR FILING DATE: 2000-10-06
;  PRIOR APPLICATION NUMBER: 60/238,372
;  PRIOR FILING DATE: 2000-10-06
;  PRIOR APPLICATION NUMBER: 60/238,383
;  PRIOR FILING DATE: 2000-10-06
;  PRIOR APPLICATION NUMBER: 60/238,382
;  PRIOR FILING DATE: 2000-10-06
;  PRIOR APPLICATION NUMBER: 60/275,892
;  PRIOR FILING DATE: 2001-03-14
;  PRIOR APPLICATION NUMBER: 60/296,860
;  PRIOR FILING DATE: 2001-06-08
;  NUMBER OF SEQ ID NOS: 198
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 55
;   LENGTH: 2995
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (64)..(2901)
US-09-972-211-55

Alignment Scores:

| | | | |
|---|---|---|---|
| Pred. No.: | 3.79e-252 | Length: | 2995 |
| Score: | 2490.50 | Matches: | 496 |
| Percent Similarity: | 67.86% | Conservative: | 150 |
| Best Local Similarity: | 52.10% | Mismatches: | 247 |
| Query Match: | 52.03% | Indels: | 59 |
| DB: | 13 | Gaps: | 18 |

US-09-970-944-2 (1-899) x US-09-972-211-55 (1-2995)

```
Qy          1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrp--- 19
               |||      |||  |||       |||||||||       :::|||      |||
Db         64 ATGGGGGCCCGGAGCGGAGCTCGGGGCGCGCTGCTGCTGGCACTGCTGCTCTGCTGGGAC 123

Qy         20 ------LeuArgGlySerGlyAlaGlnGlnSerAlaThrVal---AlaAsnProValPro 36
                    |||     :::|||  ::::::|||      |||       :::     |||
Db        124 CCGAGGCTGAGCCAAGCAGGTAGGAAGCGATCGGGTGAGGTGCTCCCTGACTCCTTCCCG 183

Qy         37 GlyAlaAsnProAspLeuLeuProHisPheLeuValGluProGluAspValTyrIleVal 56
                 |||      :::    |||||:::||||||   ||||||:::|||   |||||||||
Db        184 TCAGCGCCAGCAGAGCCGCTGCCCTACTTCCTGCAGGAGCCACAGGACGCCTACATTGTG 243

Qy         57 LysAsnLysProValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLys 76
```

```
Db    244 AAGAACAAGCCTGTGGAGCTCCGCTGCCGCGCCTTCCCCGCCACACAGATCTACTTCAAG 303
Qy     77 CysAsnGlyGluTrpValArgGlnValAspHisValIleGluArgSerThrAspGlySer 96
Db    304 TGCAACGGCGAGTGGGTCAGCCAGAACGACCACGTCACACAGGAAGGCCTGGATGAGGCC 363
Qy     97 Ser---------------GlyGluProThrMetGluValArgIleAsnValSerArgGln 111
Db    364 ACCCTGGGGGCGCGGGGCGGCCTGCGGGTGCGCGAGGTGCAGATCGAGGTGTCGCGGCAG 423
Qy    112 GlnValGluLysValPheGlyLeuGluGluTyrTrpCysGlnCysValAlaTrpSerSer 131
Db    424 CAGGTGGAGGAGCTCTTTGGGCTGGAGGATTACTGGTGCCAGTGCGTGGCCTGGAGCTCC 483
Qy    132 SerGlyThrThrLysSerGlnLysAlaTyrIleArgIleAlaArgLeuArgLysAsnPhe 151
Db    484 GCGGGCACCACCAAGAGTCGCCGAGCCTACGTCCGCATCGCCTGTCTGCGCAAGAACTTC 543
Qy    152 GluGlnGluProLeuAlaLysGluValSerLeuGluGlnGlyIleValLeuProCysArg 171
Db    544 GATCAGGAGCCTCTGGGCAAGGAGGTGCCCCTGGACCATGAGGTTCTCCTGCAGTGCCGC 603
Qy    172 ProProGluGlyIleProProAlaGluValGluTrpLeuArgAsnGluAspLeuValAsp 191
Db    604 CCGCCGGAGGGGGTGCCTGTGGCCGAGGTGGAATGGCTCAAGAATGAGGATGTCATCGAC 663
Qy    192 ProSerLeuAspProAsnValTyrIleThrArgGluHisSerLeuValValArgGlnAla 211
Db    664 CCCACCCAGGACACCAACTTCCTGCTCACCATCGACCACAACCTCATCATCCGCCAGGCC 723
Qy    212 ArgLeuAlaAspThrAlaAsnTyrThrCysValAlaLysAsnIleValAlaArgArgArg 231
Db    724 CGCCTGTCGGACACTGCCAACTATACCTGCGTGGCCAAGAACATCGTGGCCAAACGCCGG 783
Qy    232 SerAlaSerAlaAlaValIleValTyrValAsnGlyGlyTrpSerThrTrpThrGluTrp 251
Db    784 AGCACCACTGCCACCGTCATCGTCTACGTGAATGGCGGCTGGTCCAGCTGGGCAGAGTGG 843
Qy    252 SerValCysSerAlaSerCysGlyArgGlyTrpGlnLysArgSerArgSerCysThrAsn 271
Db    844 TCACCCTGCTCCAACCGCTGTGGCCGAGGCTGGCAGAAGCGCACCCGGACCTGCACCAAC 903
Qy    272 ProAlaProLeuAsnGlyGlyAlaPheCysGluGlyGlnAsnValHisAspArgThrVal 291
Db    904 CCCGCTCCACTCAACGGAGGGGCCTTCTGCGAGGGCCAG---GCATTCCAGAAGACCGCC 960
Qy    292 SerSerLeuLeuValSerValAspGlySerTrpSerProTrpSerLysTrpSerAlaCys 311
Db    961 TGCACCACCATCTGCCCAGTCGATGGGGCGTGGACGGAGTGGAGCAAGTGGTCAGCCTGC 1020
Qy    312 GlyLeuAspCysThrHisTrpArgSerArgGluCysSerAspProAlaProArgAsnGly 331
Db   1021 AGCACTGAGTGTGCCCACTGGCGTAGCCGCGAGTGCATGGCGCCCCCACCCCAGAACGGA 1080
Qy    332 GlyGluGluCysGlnGlyThrAspLeuAspThrArgAsnCysThrSerAspLeuCysVal 351
Db   1081 GGCCGTGACTGCAGCGGGACGCTGCTCGACTCTAAGAACTGCACAGATGGGCTGTGCATG 1140
Qy    352 HisSer------------------------AlaSerGlyProGluAspValAlaLeuTyr 363
Db   1141 CAAAGTGAGCCTGTCCCCGCAGTGCTGGAGGCCTCAGGG------GATGCGGCGCTGTAT 1194
Qy    364 ValGlyLeu---IleAlaValAlaValCysLeuValLeuLeuLeuLeuValLeuIleLeu 382
Db   1195 GCGGGGCTCGTGGTGGCCATCTTCGTGGTCGTGGCAATCCTCATGGCGGTGGGGGTGGTG 1254
Qy    383 ValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSer---IleLeu 401
Db   1255 GTGTACCGCCGCAACTGCCGTGACTTCGACACAGACATCACTGACTCATCTGCTGCCCTG 1314
Qy    402 ThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeuLeu 421
Db   1315 ACTGGTGGTTTCCACCCCGTCAACTTTAAGACGGCAAGGCCCAGTAACCCGCAGCTCCTA 1374
Qy    422 ------ThrIleGlnProAspLeuSerThrThrThr---ThrTyrGlnGlySerLeuCys 438
Db   1375 CACCCCTCTGTGCCTCCTGACCTGACAGCCAGCGCCGGCATCTACCGCGGACCCGTGTAT 1434
Qy    439 ProArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerPro 458
Db   1435 GCCCTGCAGGACTCC---ACCGACAAAATCCCCATGACCAACTCTCCTCTGCTGGACCCC 1491
Qy    459 LeuGlyGlyGlyArgHisThrLeuHisHisSerSerProThr------------------ 472
Db   1492 TTACCCAGCCTTAAGGTCAAGGTCTACAGCTCCAGCACCACGGGCTCTGGGCCAGGCCTG 1551
Qy    473 SerGluAlaGluGluPheValSerArgLeuSerThrGlnAsnTyr--------------- 487
Db   1552 GCAGATGGGGCTGACCTGCTGGGGGTCTTGCCGCCTGGCACATACCCTAGCGATTTCGCC 1611
Qy    488 ---------------------PheArgSer------------------------------ 490
Db   1612 CGGGACACCCACTTCCTGCACCTGCGCAGCGCCAGCCTCGGTTCCCAGCAGCTCTTGGGC 1671
Qy    491 LeuProArgGlyThrSerAsnMetThrTyrGlyThrPheAsnPheLeuGlyGlyArgLeu 510
Db   1672 CTGCCCCGAGACCCAGGGAGCAGCGTCAGCGGCACCTTTGGCTGCCTGGGTGGGAGGCTC 1731
Qy    511 MetIleProAsnThrGlyIleSerLeuLeuIleProProAspAlaIleProArgGlyLys 530
Db   1732 AGCATCCCCGGCACAGGTGTCAGCTTGCTGGTGCCCAATGGAGCCATTCCCCAGGGCAAG 1791
Qy    531 IleTyrGluIleTyrLeuThrLeuHisLysProGluAspValArgLeuProLeuAla--- 549
Db   1792 TTCTACGAGATGTATCTACTCATCAACAAGGCAGAAAGTACC---CTGCCGCTTTCAGAA 1848
Qy    550 GlyCysGlnThrLeuLeuSerProIleValSerCysGlyProProGlyValLeuLeuThr 569
Db   1849 GGGACCCAGACAGTATTGAGCCCCTCGGTGACCTGTGGACCCACAGGCCTCCTGCTGTGC 1908
Qy    570 ArgProValIleLeuAlaMetAspHisCysGlyGluProSerProAspSerTrpSerLeu 589
Db   1909 CGCCCCGTCATCCTCACCATGCCCCACTGTGCCGAAGTCAGTGCCCGTGACTGGATCTTT 1968
Qy    590 ArgLeuLysLysGlnSerCysGluGlySerTrpGluGlnAspValLeuHisLeuGlyGlu 609
Db   1969 CAGCTCAAGACCCAGGCCCACCAGGGCCACTGGGAGCAGGAGGTGGTGACCCTGGATGAG 2028
Qy    610 GluAlaProSerHisLeuTyrTyrCysGlnLeuGluAlaSerAlaCysTyrValPheThr 629
Db   2029 GAGACCCTGAACACACCCTGCTACTGCCAGCTGGAGCCCAGGGCCTGTCACATCCTGCTG 2088
Qy    630 GluGlnLeuGlyArgPheAlaLeuValGlyGluAlaLeuSerValAlaAlaAlaLysArg 649
Db   2089 GACCAGCTGGGCACCTACGTGTTCACGGGCGAGTCCTATTCCCGCTCAGCAGTCAAGCGG 2148
Qy    650 LeuLysLeuLeuLeuPheAlaProValAlaCysThrSerLeuGluTyrAsnIleArgVal 669
Db   2149 CTCCAGCTGGCCGTCTTCGCCCCCGCCCTCTGCACCTCCCTGGAGTACAGCCTCCGGGTC 2208
Qy    670 TyrCysLeuHisAspThrHisAspAlaLeuLysGluValValGlnLeuGluLysGlnLeu 689
Db   2209 TACTGCCTGGAGGACACGCCTGTAGCACTGAAGGAGGTGCTGGAGCTGGAGCGGACTCTG 2268
Qy    690 GlyGlyGlnLeuIleGlnGluProArgValLeuHisPheLysAspSerTyrHisAsnLeu 709
Db   2269 GGCGGATACTTGGTGGAGGAGCCGAAACCGCTAATGTTCAAGGACAGTTACCACAACCTG 2328
Qy    710 ArgLeuSerIleHisAspValProSerSerLeuTrpLysSerLysLeuLeuValSerTyr 729
Db   2329 CGCCTCTCCCTCCATGACCTCCCCCATGCCCATTGGAGGAGCAAGCTGCTGGCCAAATAC 2388
Qy    730 GlnGluIleProPheTyrHisIleTrpAsnGlyThrGlnArgTyrLeuHisCysThrPhe 749
Db   2389 CAGGAGATCCCCTTCTATCACATTTGGAGTGGCAGCCAGAAGGCCCTCCACTGCACTTTC 2448
```

```
Qy        750 ThrLeuGluArgValSerProSerThrSerAspLeuAlaCysLysLeuTrpValTrpGln 769
              |||||||||||   |||   :::::::::::|||  ||||||:::  |||  |||
Db       2449 ACCCTGGAGAGGCACAGCTTGGCCTCCACAGAGCTCACCTGCAAGATCTGCGTGCGGCAA 2508

Qy        770 ValGluGlyAspGlyGlnSerPheSerIleAsnPheAsnIleThrLysAspThrArgPhe 789
              ||||||||:::||||||   |||   ::::::      :::      :::|||
Db       2509 GTGGAAGGGGAGGGCCAGATATTCCAGCTGCATACCACTCTGGCA---GAGACACCTGCT 2565

Qy        790 AlaGluLeuLeuAlaLeuGluSerGluAlaGly------ValProAlaLeuValGlyPro 807
                    |||    |||   |||     |||      |||        :::||||||
Db       2566 GGCTCCCTGGACACTCTCTGCTCTGCCCCTGGCAGCACTGTCACCACCCAGCTGGGACCT 2625

Qy        808 SerAlaPheLysIleProPheLeuIleArgGlnLysIleIleSerSerLeuAspProPro 827
                 ||||||||||||||        ||||||||||||||    :::|||||||||   |||
Db       2626 TATGCCTTCAAGATCCCACTGTCCATCCGCCAGAAGATATGCAACAGCCTAGATGCCCCC 2685

Qy        828 CysArgArgGlyAlaAspTrpArgThrLeuAlaGlnLysLeuHisLeuAspSerHisLeu 847
                    ||||||   |||||||||   ||||||||||||||||    :::|||   :::|||
Db       2686 AACTCACGGGGCAATGACTGGCGGATGTTAGCACAGAAGCTCTCTATGGACCGGTACCTG 2745

Qy        848 SerPhePheAlaSerLysProSerProThrAlaMetIleLeuAsnLeuTrpGluAlaArg 867
              ::::::||||||:::|||   |||||||||   :::||||||:::||||||||||||
Db       2746 AATTACTTTGCCACCAAAGCGAGCCCCACGGGTGTGATCCTGGACCTCTGGGAAGCTCTG 2805

Qy        868 HisPheProAsnGlyAsnLeuSerGlnLeuAlaAlaAlaValAlaGlyLeuGlyGlnPro 887
                       :::|||:::|||:::   ||||||:::|||:::      :::|||:::
Db       2806 CAGCAGGACGATGGGGACCTCAACAGCCTGGCGAGTGCCTTGGAGGAGATGGGCAAGAGT 2865

Qy        888 AspAlaGlyLeuPheThrValSerGluAlaGluCys 899
              :::      :::         ::::::   :::|||
Db       2866 GAGATGCTGGTGGCTGTGGCCACCGACGGGGACTGC 2901
```

RESULT 11
US-10-096-625-55
; Sequence 55, Application US/10096625
; Publication No. US20040068095A1
; GENERAL INFORMATION:
; APPLICANT: Shimkets, Richard A
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Mezes, Peter S
; APPLICANT: Rastelli, Luca
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Grosse, William M
; APPLICANT: Alsobrook II, John P
; APPLICANT: Lepley, Denise M
; APPLICANT: Spytek, Kimberly Ann
; APPLICANT: Li, Li
; APPLICANT: Edinger, Shlomit
; APPLICANT: Gerlach, Valerie
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John R
; APPLICANT: Gunther, Erik
; APPLICANT: Millet, Isabelle
; APPLICANT: Stone, David J
; APPLICANT: Smithson, Glennda
; APPLICANT: Szekeres Jr, Edward S
; APPLICANT: Ji, Weizhen
; TITLE OF INVENTION: No. US20040068095A1el Human Proteins, Polynucleotides Encoding Th
; TITLE OF INVENTION: Methods Of Using The Same
; FILE REFERENCE: 21402-141 CIP
; CURRENT APPLICATION NUMBER: US/10/096,625
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 09/972,211
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 60/238,325
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/238,323
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/238,400
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,397
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,401
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,379
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,402
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 30/238,384
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,373
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,372
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,383
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,382
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/275,892
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/296,860
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 200
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 55
; LENGTH: 2995
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (64)..(2901)
US-10-096-625-55

```
Alignment Scores:
Pred. No.:              3.79e-252     Length:         2995
Score:                  2490.50       Matches:        496
Percent Similarity:     67.86%        Conservative:   150
Best Local Similarity:  52.10%        Mismatches:     247
Query Match:            52.03%        Indels:         59
DB:                     13            Gaps:           18
```

US-09-970-944-2 (1-899) x US-10-096-625-55 (1-2995)

```
Qy          1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrp--- 19
              |||      |||   |||         |||||||||      :::|||      |||
Db         64 ATGGGGGCCCGGAGCGGAGCTCGGGGCGCGCTGCTGCTGGCACTGCTCTGCTGGGAC 123

Qy         20 ------LeuArgGlySerGlyAlaGlnGlnSerAlaThrVal---AlaAsnProValPro 36
                    |||      :::|||   ::::::|||      |||      :::      |||
Db        124 CCGAGGCTGAGCCAAGCAGGTAGGAAGCGATCGGGTGAGGTGCTCCCTGACTCCTTCCCG 183

Qy         37 GlyAlaAsnProAspLeuLeuProHisPheLeuValGluProGluAspValTyrIleVal 56
                 |||      :::   ||||||:::||||||   ||||||:::|||   ||||||||
Db        184 TCAGCGCCAGCAGAGCCGCTGCCCTACTTCCTGCAGGAGCCACAGGACGCCTACATTGTG 243

Qy         57 LysAsnLysProValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLys 76
              ||||||||||||||   |||   |||:::|||   |||||||||||||||:::||||||
Db        244 AAGAACAAGCCTGTGGAGCTCCGCTGCCGCGCCTTCCCCGCCACACAGATCTACTTCAAG 303

Qy         77 CysAsnGlyGluTrpValArgGlnValAspHisValIleGluArgSerThrAspGlySer 96
              ||||||||||||||||||   |||   |||||||||   :::         |||   :::
Db        304 TGCAACGGCGAGTGGGTCAGCCAGAACGACCACGTCACACAGGAAGGCCTGGATGAGGCC 363

Qy         97 Ser---------------GlyGluProThrMetGluValArgIleAsnValSerArgGln 111
              :::               |||         ||||||:::|||   ||||||||||||
Db        364 ACCCTGGGGGCGCGGGGCGGCCTGCGGGTGCGCGAGGTGCAGATCGAGGTGTCGCGGCAG 423

Qy        112 GlnValGluLysValPheGlyLeuGluGluTyrTrpCysGlnCysValAlaTrpSerSer 131
              |||||||||::::::||||||||||||:::||||||||||||||||||||||||||||||
Db        424 CAGGTGGAGGAGCTCTTTGGGCTGGAGGATTACTGGTGCCAGTGCGTGGCCTGGAGCTCC 483

Qy        132 SerGlyThrThrLysSerGlnLysAlaTyrIleArgIleAlaArgLeuArgLysAsnPhe 151
              :::||||||||||||||||||::::::||||||:::|||||||||   ||||||||||||
```

```
Db     484 GCGGGCACCACCAAGAGTCGCCGAGCCTACGTCCGCATCGCCTGTCTGCGCAAGAACTTC 543
Qy     152 GluGlnGluProLeuAlaLysGluValSerLeuGluGlnGlyIleValLeuProCysArg 171
Db     544 GATCAGGAGCCTCTGGGCAAGGAGGTGCCCCTGGACCATGAGGTTCTCCTGCAGTGCCGC 603
Qy     172 ProProGluGlyIleProProAlaGluValGluTrpLeuArgAsnGluAspLeuValAsp 191
Db     604 CCGCCGGAGGGGGTGCCTGTGGCCGAGGTGGAATGGCTCAAGAATGAGGATGTCATCGAC 663
Qy     192 ProSerLeuAspProAsnValTyrIleThrArgGluHisSerLeuValValArgGlnAla 211
Db     664 CCCACCCAGGACACCAACTTCCTGCTCACCATCGACCACAACCTCATCATCCGCCAGGCC 723
Qy     212 ArgLeuAlaAspThrAlaAsnTyrThrCysValAlaLysAsnIleValAlaArgArgArg 231
Db     724 CGCCTGTCGGACACTGCCAACTATACCTGCGTGGCCAAGAACATCGTGGCCAAACGCCGG 783
Qy     232 SerAlaSerAlaAlaValIleValTyrValAsnGlyGlyTrpSerThrTrpThrGluTrp 251
Db     784 AGCACCACTGCCACCGTCATCGTCTACGTGAATGGCGGCTGGTCCAGCTGGGCAGAGTGG 843
Qy     252 SerValCysSerAlaSerCysGlyArgGlyTrpGlnLysArgSerArgSerCysThrAsn 271
Db     844 TCACCCTGCTCCAACCGCTGTGGCCGAGGCTGGCAGAAGCGCACCCGGACCTGCACCAAC 903
Qy     272 ProAlaProLeuAsnGlyGlyAlaPheCysGluGlyGlnAsnValHisAspArgThrVal 291
Db     904 CCCGCTCCACTCAACGGAGGGGCCTTCTGCGAGGGCCAG---GCATTCCAGAAGACCGCC 960
Qy     292 SerSerLeuLeuValSerValAspGlySerTrpSerProTrpSerLysTrpSerAlaCys 311
Db     961 TGCACCACCATCTGCCCAGTCGATGGGGCGTGGACGGAGTGGAGCAAGTGGTCAGCCTGC 1020
Qy     312 GlyLeuAspCysThrHisTrpArgSerArgGluCysSerAspProAlaProArgAsnGly 331
Db    1021 AGCACTGAGTGTGCCCACTGGCGTAGCCGCGAGTGCATGGCGCCCCCACCCCAGAACGGA 1080
Qy     332 GlyGluGluCysGlnGlyThrAspLeuAspThrArgAsnCysThrSerAspLeuCysVal 351
Db    1081 GGCCGTGACTGCAGCGGGACGCTGCTCGACTCTAAGAACTGCACAGATGGGCTGTGCATG 1140
Qy     352 HisSer-----------------------AlaSerGlyProGluAspValAlaLeuTyr 363
Db    1141 CAAAGTGAGCCTGTCCCCGCAGTGCTGGAGGCCTCAGGG------GATGCGGCGCTGTAT 1194
Qy     364 ValGlyLeu---IleAlaValAlaValCysLeuValLeuLeuLeuLeuValLeuIleLeu 382
Db    1195 GCGGGGCTCGTGGTGGCCATCTTCGTGGTCGTGGCAATCCTCATGGCGGTGGGGGTGGTG 1254
Qy     383 ValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSer---IleLeu 401
Db    1255 GTGTACCGCCGCAACTGCCGTGACTTCGACACAGACATCACTGACTCATCTGCTGCCCTG 1314
Qy     402 ThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeuLeu 421
Db    1315 ACTGGTGGTTTCCACCCCGTCAACTTTAAGACGGCAAGGCCCAGTAACCCGCAGCTCCTA 1374
Qy     422 ------ThrIleGlnProAspLeuSerThrThrThr---ThrTyrGlnGlySerLeuCys 438
Db    1375 CACCCCTCTGTGCCTCCTGACCTGACAGCCAGCGCCGGCATCTACCGCGGACCCGTGTAT 1434
Qy     439 ProArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerPro 458
Db    1435 GCCCTGCAGGACTCC---ACCGACAAAATCCCCATGACCAACTCTCCTCTGCTGGACCCC 1491
Qy     459 LeuGlyGlyGlyArgHisThrLeuHisHisSerSerProThr------------------ 472
Db    1492 TTACCCAGCCTTAAGGTCAAGGTCTACAGCTCCAGCACCACGGGCTCTGGGCCAGGCCTG 1551
Qy     473 SerGluAlaGluGluPheValSerArgLeuSerThrGlnAsnTyr--------------- 487
Db    1552 GCAGATGGGGCTGACCTGCTGGGGGTCTTGCCGCCTGGCACATACCCTAGCGATTTCGCC 1611

Qy     488 --------------------PheArgSer----------------------------- 490
Db    1612 CGGGACACCCACTTCCTGCACCTGCGCAGCGCCAGCCTCGGTTCCCAGCAGCTCTTGGGC 1671
Qy     491 LeuProArgGlyThrSerAsnMetThrTyrGlyThrPheAsnPheLeuGlyGlyArgLeu 510
Db    1672 CTGCCCCGAGACCCAGGGAGCAGCGTCAGCGGCACCTTTGGCTGCCTGGGTGGGAGGCTC 1731
Qy     511 MetIleProAsnThrGlyIleSerLeuLeuIleProProAspAlaIleProArgGlyLys 530
Db    1732 AGCATCCCCGGCACAGGTGTCAGCTTGCTGGTGCCCAATGGAGCCATTCCCCAGGGCAAG 1791
Qy     531 IleTyrGluIleTyrLeuThrLeuHisLysProGluAspValArgLeuProLeuAla--- 549
Db    1792 TTCTACGAGATGTATCTACTCATCAACAAGGCAGAAAGTACC---CTGCCGCTTTCAGAA 1848
Qy     550 GlyCysGlnThrLeuLeuSerProIleValSerCysGlyProProGlyValLeuLeuThr 569
Db    1849 GGGACCCAGACAGTATTGAGCCCCTCGGTGACCTGTGGACCCACAGGCCTCCTGCTGTGC 1908
Qy     570 ArgProValIleLeuAlaMetAspHisCysGlyGluProSerProAspSerTrpSerLeu 589
Db    1909 CGCCCCGTCATCCTCACCATGCCCCACTGTGCCGAAGTCAGTGCCCGTGACTGGATCTTT 1968
Qy     590 ArgLeuLysLysGlnSerCysGluGlySerTrpGluGlnAspValLeuHisLeuGlyGlu 609
Db    1969 CAGCTCAAGACCCAGGCCCACCAGGGCCACTGGGAGCAGGAGGTGGTGACCCTGGATGAG 2028
Qy     610 GluAlaProSerHisLeuTyrTyrCysGlnLeuGluAlaSerAlaCysTyrValPheThr 629
Db    2029 GAGACCCTGAACACACCCTGCTACTGCCAGCTGGAGCCCAGGGCCTGTCACATCCTGCTG 2088
Qy     630 GluGlnLeuGlyArgPheAlaLeuValGlyGluAlaLeuSerValAlaAlaAlaLysArg 649
Db    2089 GACCAGCTGGGCACCTACGTGTTCACGGGCGAGTCCTATTCCCGCTCAGCAGTCAAGCGG 2148
Qy     650 LeuLysLeuLeuLeuPheAlaProValAlaCysThrSerLeuGluTyrAsnIleArgVal 669
Db    2149 CTCCAGCTGGCCGTCTTCGCCCCCGCCCTCTGCACCTCCCTGGAGTACAGCCTCCGGGTC 2208
Qy     670 TyrCysLeuHisAspThrHisAspAlaLeuLysGluValValGlnLeuGluLysGlnLeu 689
Db    2209 TACTGCCTGGAGGACACGCCTGTAGCACTGAAGGAGGTGCTGGAGCTGGAGCGGACTCTG 2268
Qy     690 GlyGlyGlnLeuIleGlnGluProArgValLeuHisPheLysAspSerTyrHisAsnLeu 709
Db    2269 GGCGGATACTTGGTGGAGGAGCCGAAACCGCTAATGTTCAAGGACAGTTACCACAACCTG 2328
Qy     710 ArgLeuSerIleHisAspValProSerSerLeuTrpLysSerLysLeuLeuValSerTyr 729
Db    2329 CGCCTCTCCCTCCATGACCTCCCCCATGCCCATTGGAGGAGCAAGCTGCTGGCCAAATAC 2388
Qy     730 GlnGluIleProPheTyrHisIleTrpAsnGlyThrGlnArgTyrLeuHisCysThrPhe 749
Db    2389 CAGGAGATCCCCTTCTATCACATTTGGAGTGGCAGCCAGAAGGCCCTCCACTGCACTTTC 2448
Qy     750 ThrLeuGluArgValSerProSerThrSerAspLeuAlaCysLysLeuTrpValTrpGln 769
Db    2449 ACCCTGGAGAGGCACAGCTTGGCCTCCACAGAGCTCACCTGCAAGATCTGCGTGCGGCAA 2508
Qy     770 ValGluGlyAspGlyGlnSerPheSerIleAsnPheAsnIleThrLysAspThrArgPhe 789
Db    2509 GTGGAAGGGGAGGGCCAGATATTCCAGCTGCATACCACTCTGGCA---GAGACACCTGCT 2565
Qy     790 AlaGluLeuLeuAlaLeuGluSerGluAlaGly------ValProAlaLeuValGlyPro 807
Db    2566 GGCTCCCTGGACACTCTCTGCTCTGCCCCTGGCAGCACTGTCACCACCCAGCTGGGACCT 2625
Qy     808 SerAlaPheLysIleProPheLeuIleArgGlnLysIleIleSerSerLeuAspProPro 827
Db    2626 TATGCCTTCAAGATCCCACTGTCCATCCGCCAGAAGATATGCAACAGCCTAGATGCCCCC 2685
```

```
Qy        828 CysArgArgGlyAlaAspTrpArgThrLeuAlaGlnLysLeuHisLeuAspSerHisLeu 847
                      |||||    |||||||||   ||||||||||||||    :::|||    :::|||
Db       2686 AACTCACGGGGCAATGACTGGCGGATGTTAGCACAGAAGCTCTCTATGGACCGGTACCTG 2745

Qy        848 SerPhePheAlaSerLysProSerProThrAlaMetIleLeuAsnLeuTrpGluAlaArg 867
              ::::::||||||:::|||    |||||||||    :::||||||:::|||||||||||||
Db       2746 AATTACTTTGCCACCAAAGCGAGCCCCACGGGTGTGATCCTGGACCTCTGGGAAGCTCTG 2805

Qy        868 HisPheProAsnGlyAsnLeuSerGlnLeuAlaAlaAlaValAlaGlyLeuGlyGlnPro 887
                          :::|||:::|||:::   ||||||:::|||:::         :::|||:::
Db       2806 CAGCAGGACGATGGGGACCTCAACAGCCTGGCGAGTGCCTTGGAGGAGATGGGCAAGAGT 2865

Qy        888 AspAlaGlyLeuPheThrValSerGluAlaGluCys 899
              :::      :::         ::::::   :::|||
Db       2866 GAGATGCTGGTGGCTGTGGCCACCGACGGGGACTGC 2901
```

RESULT 12
US-10-087-684-1
; Sequence 1, Application US/10087684
; Publication No. US20040029116A1
; GENERAL INFORMATION:
;  APPLICANT: Edinger, Shlomit R.
;  APPLICANT:  MacDougall, John R.
;  APPLICANT:  Millet, Isabelle
;  APPLICANT:  Ellerman, Karen
;  APPLICANT:  Stone, David J.
;  APPLICANT:  Grosse, William M.
;  APPLICANT:  Lepley, Denise M.
;  APPLICANT:  Rieger, Daniel K.
;  APPLICANT:  Burgess, Cathereine E.
;  APPLICANT:  Casman, Stacie, J.
;  APPLICANT:  Spytek, Kimberly A.
;  APPLICANT:  Boldog, Ferenc L.
;  APPLICANT:  Li, Li
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Mishra, Vishnu
;  APPLICANT:  Shenoy, Suresh G.
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Tchernev, Velizar T.
;  APPLICANT:  Vernet, Corine A.M.
;  APPLICANT:  Zerhusen, Bryan D.
;  APPLICANT:  Malyankar, Uriel M.
;  APPLICANT:  Guo, Xiaojia
;  APPLICANT:  Miller, Charles E.
;  APPLICANT:  Gangolli, Esha A.
;  TITLE OF INVENTION: PROTEINS AND NUCLEIC ACIDS ENCODING SAME
;  FILE REFERENCE: 21402-214 CIP
;  CURRENT APPLICATION NUMBER: US/10/087,684
;  CURRENT FILING DATE:  2003-03-10
;  PRIOR APPLICATION NUMBER: 60/253,834
;  PRIOR FILING DATE: 2000-11-29
;  PRIOR APPLICATION NUMBER: 60/250,926
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: 60/264,180
;  PRIOR FILING DATE: 2001-01-25
;  PRIOR APPLICATION NUMBER: 60/274,194
;  PRIOR FILING DATE: 2001-03-08
;  PRIOR APPLICATION NUMBER: 60/313,656
;  PRIOR FILING DATE: 2001-08-20
;  PRIOR APPLICATION NUMBER: 60/327,456
;  PRIOR FILING DATE: 2001-10-05
;  NUMBER OF SEQ ID NOS: 220
;  SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 1
;   LENGTH: 2860
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (59)..(2857)
US-10-087-684-1

Alignment Scores:

```
Pred. No.:               1.72e-251     Length:         2860
Score:                   2484.00       Matches:        493
Percent Similarity:      68.30%        Conservative:   149
Best Local Similarity:   52.45%        Mismatches:     250
Query Match:             51.89%        Indels:         48
DB:                      13            Gaps:           18

US-09-970-944-2 (1-899) x US-10-087-684-1 (1-2860)

Qy          1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrp--- 19
              |||      |||   |||         |||||||||      :::|||      |||
Db         59 ATGGGGGCCCGGAGCGGAGCTCGGGGCGCGCTGCTGCTGGCACTGCTGCTCTGCTGGGAC 118

Qy         20 ------LeuArgGlySerGlyAlaGlnGln---SerAlaThrValAlaAsnProValPro 36
                    |||      :::|||               |||   :::   :::      |||
Db        119 CCGAGGCTGAGCCAAGCAGGCACTGATTCTGGCAGCGAGGTGCTCCCTGACTCCTTCCCG 178

Qy         37 GlyAlaAsnProAspLeuLeuProHisPheLeuValGluProGluAspValTyrIleVal 56
                 |||      :::   ||||||:::||||||   ||||||:::|||   |||||||||
Db        179 TCAGCGCCAGCAGAGCCGCTGCCCTACTTCCTGCAGGAGCCACAGGACGCCTACATTGTG 238

Qy         57 LysAsnLysProValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLys 76
              |||||||||||||||   |||   |||:::|||   ||||||||||||||||:::|||||
Db        239 AAGAACAAGCCTGTGGAGCTCCGCTGCCGCGCCTTCCCCGCCACACAGATCTACTTCAAG 298

Qy         77 CysAsnGlyGluTrpValArgGlnValAspHisValIleGluArgSerThrAspGlySer 96
              ||||||||||||||||||   |||   |||||||||   :::         |||   :::
Db        299 TGCAACGGCGAGTGGGTCAGCCAGAACGACCACGTCACACAGGAAGGCCTGGATGAGGCC 358

Qy         97 SerGlyGluProThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysVal 116
              :::|||            ||||||:::|||   |||||||||||||||||||||::::::
Db        359 ACCGGTCTGCGGGTGCGCGAGGTGCAGATCGAGGTGTCGCGGCAGCAGGTGGAGGAGCTC 418

Qy        117 PheGlyLeuGluGluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLys 136
              ||||||||||||:::|||||||||||||||||||||||||||||||:::|||||||||||
Db        419 TTTGGGCTGGAGGATTACTGGTGCCAGTGCGTGGCCTGGAGCTCCGCGGGCACCACCAAG 478

Qy        137 SerGlnLysAlaTyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeu 156
              |||::::::||||||:::|||||||||   ||||||||||||:::|||||||||||||||
Db        479 AGTCGCCGAGCCTACGTCCGCATCGCCTACCTGCGCAAGAACTTCGATCAGGAGCCTCTG 538

Qy        157 AlaLysGluValSerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIle 176
                 |||||||||   |||:::      ::::::|||   |||||||||||||||||::::::
Db        539 GGCAAGGAGGTGCCCCTGGACCATGAGGTTCTCCTGCAGTGCCGCCCGCCGGAGGGGGTG 598

Qy        177 ProProAlaGluValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspPro 196
              |||   ||||||||||||||||||:::|||||||||::::::||||||:::   |||
Db        599 CCTGTGGCCGAGGTGGAATGGCTCAAGAATGAGGATGTCATCGACCCCACCCAGGACACC 658

Qy        197 AsnValTyrIleThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThr 216
              |||      :::|||   :::|||:::|||::::::||||||||||||||||:::||||||
Db        659 AACTTCCTGCTCACCATCGACCACAACCTCATCATCCGCCAGGCCCGCCTGTCGGACACT 718

Qy        217 AlaAsnTyrThrCysValAlaLysAsnIleValAlaArgArgArgSerAlaSerAlaAla 236
              ||||||||||||||||||||||||||||||||||||:::||||||         :::|||
Db        719 GCCAACTATACCTGCGTGGCCAAGAACATCGTGGCCAAACGCCGGAGCACCACTGCCACC 778

Qy        237 ValIleValTyrValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAla 256
              |||||||||||||||||||||||||||:::|||   |||||||||   ||||||
Db        779 GTCATCGTCTACGTGAATGGCGGCTGGTCCAGCTGGGCAGAGTGGTCACCCTGCTCCAAC 838

Qy        257 SerCysGlyArgGlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsn 276
                 ||||||||||||||||||||||:::|||:::||||||||||||||||||||||||
Db        839 CGCTGTGGCCGAGGCTGGCAGAAGCGCACCCGGACCTGCACCAACCCCGCTCCACTCAAC 898

Qy        277 GlyGlyAlaPheCysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuVal 296
              ||||||||||||||||||||||                  :::|||      :::   :::
Db        899 GGAGGGGCCTTCTGCGAGGGCCAG---GCATTCCAGAAGACCGCCTGCACCACCATCTGC 955

Qy        297 SerValAspGlySerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThr 316
                 |||||||||:::|||:::   ||||||||||||||||||||         :::|||
Db        956 CCAGTCGATGGGGCGTGGACGGAGTGGAGCAAGTGGTCAGCCTGCAGCACTGAGTGTGCC 1015
```

```
Qy    317 HisTrpArgSerArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGln 336
          ||||||||||||||||||||      |||   |||:::||||||||    :::|||
Db   1016 CACTGGCGTAGCCGCGAGTGCATGGCGCCCCCACCCCAGAACGGAGGCCGTGACTGCAGC 1075

Qy    337 GlyThrAspLeuAspThrArgAsnCysThrSerAspLeuCysValHis---SerAlaSer 355
          ||||||  ||||||::::::||||||||         ||||||:::         ||||||
Db   1076 GGGACGCTGCTCGACTCTAAGAACTGCACAGATGGCTGTGCATGCAACTGGAGGCCTCA 1135

Qy    356 GlyProGluAspValAlaLeuTyrValGlyLeu---IleAlaValAlaValCysLeuVal 374
          |||      |||  |||||||||   ||||||   :::|||:::   |||    :::
Db   1136 GGG------GATGCGGCGCTGTATGCGGGGCTCGTGGTGGCCATCTTCGTGGTCGTGGCA 1189

Qy    375 LeuLeuLeuLeuValLeuIleLeuValTyrCysArgLysLysGluGlyLeuAspSerAsp 394
          :::|||:::    |||    ::::::||||||  |||              |||:::|||
Db   1190 ATCCTCATGGCGGTGGGGGTGGTGGTGTACCGCCGCAACTGCCGTGACTTCGACACAGAC 1249

Qy    395 ValAlaAspSerSer---IleLeuThrSerGlyPheGlnProValSerIleLysProSer 413
          :::   |||||||||      ||||||   ||||||   ||||||:::   |||    :::
Db   1250 ATCACTGACTCATCTGCTGCCCTGACTGGTGGTTTCCACCCCGTCAACTTTAAGACGGCA 1309

Qy    414 LysAlaAspAsnProHisLeuLeu------ThrIleGlnProAspLeuSerThrThrThr 431
          :::         ||||||   ||||||        ::::::   |||||||||:::    :::
Db   1310 AGGCCCAGTAACCCGCAGCTCCTACACCCCTCTGTGCCTCCTGACCTGACAGCCAGCGCC 1369

Qy    432 ---ThrTyrGlnGlySerLeuCysProArgGlnAspGlyProSerProLysPheGlnLeu 450
             |||:::|||      :::         ||||||      :::   |||      :::
Db   1370 GGCATCTACCGCGGACCCGTGTATGCCCTGCAGGACTCC---ACCGACAAAATCCCCATG 1426

Qy    451 ThrAsnGlyHisLeuLeuSerProLeuGlyGlyGlyArgHisThrLeuHisHisSerSer 470
          ||||||      ||||||   ||||||         :::      ::::::   ||||||
Db   1427 ACCAACTCTCCTCTGCTGGACCCCTTACCCAGCCTTAAGGTCAAGGTCTACAGCTCCAGC 1486

Qy    471 ProThr------------------SerGluAlaGluGluPheValSerArgLeuSerThr 484
            |||                   ::::::   :::   :::      |||
Db   1487 ACCACGGGCTCTGGGCCAGGCCTGGCAGATGGGGCTGACCTGCTGGGGGTCTTGCCGCCT 1546

Qy    485 GlnAsnTyr------------------------------------PheArgSer------ 490
             |||                                         ||||||
Db   1547 GGCACATACCCTAGCGATTTCGCCCGGGACACCCACTTCCTGCACCTGCGCAGCGCCAGC 1606

Qy    491 ------------------------LeuProArgGlyThrSerAsnMetThrTyrGlyThr 502
                                  |||||||||      :::         ||||||
Db   1607 CTCGGTTCCCAGCAGCTCTTGGGCCTGCCCCGAGACCCAGGGAGCAGCGTCAGCGGCACC 1666

Qy    503 PheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeuLeuIlePro 522
          |||      |||||||||||||||   ||||||   ||||||:::||||||||:::|||
Db   1667 TTTGGCTGCCTGGGTGGAGGCTCAGCATCCCCGGCACAGGGGTCAGCTTGCTGGTGCCC 1726

Qy    523 ProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHisLysProGlu 542
                |||||||||:::||||||   ||||||:::||||||   ::::::|||   |||
Db   1727 AATGGAGCCATTCCCCAGGGCAAGTTCTACGAGATGTATCTACTCATCAACAAGGCAGAA 1786

Qy    543 AspValArgLeuProLeuAla---GlyCysGlnThrLeuLeuSerProIleValSerCys 561
                |||||||||:::   |||   ||||||:::||||||||    |||:::|||
Db   1787 AGTACC---CTGCCGCTTTCAGAAGGGACCCAGACAGTATTGAGCCCCTCGGTGACCTGT 1843

Qy    562 GlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGlyGlu 581
          ||||||   |||:::||||||   ||||||||||||||   |||   ||||||   |||
Db   1844 GGACCCACAGGCCTCCTGCTGTGCCGCCCCGTCATCCTCACCATGCCCCACTGTGCCGAA 1903

Qy    582 ProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySerTrpGlu 601
            |||      |||      :::||||||   |||:::   :::|||   ||||||
Db   1904 GTCAGTGCCCGTGACTGGATCTTTCAGCTCAAGACCCAGGCCCACCAGGGCCACTGGGAG 1963

Qy    602 GlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGlnLeuGlu 621
             :::|||:::   |||   ||||||      :::         |||||||||||||||
Db   1964 ---GAGGTGGTGACCCTGGATGAGGAGACCCTGAACACACCCTGCTACTGCCAGCTGGAG 2020

Qy    622 AlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGluAla 641
                ||||||::::::      :::|||||||||   :::         ||||||:::
Db   2021 CCCAGGGCCTGTCACATCCTGCTGGACCAGCTGGGCACCTACGTGTTCACGGGCGAGTCC 2080

Qy    642 LeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAlaCysThr 661
          |||    :::|||   ||||||||:::|||    :::||||||||    ||||||
Db   2081 TATTCCCGCTCAGCAGTCAAGCGGCTCCAGCTGGCCGTCTTCGCCCCCGCCCTCTGCACC 2140

Qy    662 SerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuLysGlu 681
          ||||||||||||::::::||||||||||||||   ||||||      ||||||||||||
Db   2141 TCCCTGGAGTACAGCCTCCGGGTCTACTGCCTGGAGGACACGCCTGTAGCACTGAAGGAG 2200

Qy    682 ValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgValLeuHis 701
          |||::::::||||||:::   |||||||||   |||::::::||||||:::   |||
Db   2201 GTGCTGGAGCTGGAGCGGACTCTGGGCGGATACTTGGTGGAGGAGCCGAAACCGCTAATG 2260

Qy    702 PheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSerLeuTrp 721
          |||||||||||||||||||||||||||||:::||||||:::|||   :::   |||
Db   2261 TTCAAGGACAGTTACCACAACCTGCGCCTCTCCCTCCATGACCTCCCCCATGCCCATTGG 2320

Qy    722 LysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsnGlyThr 741
          :::||||||||||||      |||||||||||||||||||||||||||:::|||:::
Db   2321 AGGAGCAAGCTGCTGGCCAAATACCAGGAGATCCCCTTCTATCACATTTGGAGTGGCAGC 2380

Qy    742 GlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSerAspLeu 761
          |||:::   ||||||||||||||||||||||||   |||   ::::::::::::|||
Db   2381 CAGAAGGCCCTCCACTGCACTTTCACCCTGGAGAGGCACAGCTTGGCCTCCACAGAGCTC 2440

Qy    762 AlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIleAsnPhe 781
             ||||||:::   |||   ||||||||||||:::||||||   |||   ::::::
Db   2441 ACCTGCAAGATCTGCGTGCGGCAAGTGGAAGGGGAGGGCCAGATATTCCAGCTGCATACC 2500

Qy    782 AsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAlaGly--- 800
                                 |||   :::      |||   |||      |||
Db   2501 ACTCTGGCA---GAGACACCTGCTGGCTCCCTGGACACTCTCTGCTCTGCCCCTGGCAGC 2557

Qy    801 ---ValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
             |||      :::||||||   ||||||||||||||      ||||||||||||
Db   2558 ACTGTCACCACCCAGCTGGGACCTTATGCCTTCAAGATCCCACTGTCCATCCGCCAGAAG 2617

Qy    820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839
          |||   :::|||||||||   |||      ||||||   |||||||||   |||||||||
Db   2618 ATATGCAACAGCCTAGATGCCCCCAACTCACGGGGCAATGACTGGCGGATGTTAGCACAG 2677

Qy    840 LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859
          ||||||   :::|||   :::|||::::::||||||:::|||   |||||||||   :::
Db   2678 AAGCTCTCTATGGACCGGTACCTGAATTACTTTGCCACCAAAGCGAGCCCCACGGGTGTG 2737

Qy    860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879
          ||||||:::|||||||||||||               :::|||:::|||:::   ||||||:::
Db   2738 ATCCTGGACCTCTGGGAAGCTCTGCAGCAGGACGATGGGGACCTCAACAGCCTGGCGAGT 2797

Qy    880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
          |||:::      :::|||:::   :::         :::         ::::::   :::|||
Db   2798 GCCTTGGAGGAGATGGGCAAGAGTGAGATGCTGGTGGCTGTGGCCACCGACGGGGACTGC 2857
```

RESULT 13
US-10-087-684-3
; Sequence 3, Application US/10087684
; Publication No. US20040029116A1
; GENERAL INFORMATION:
;  APPLICANT: Edinger, Shlomit R.
;  APPLICANT: MacDougall, John R.
;  APPLICANT: Millet, Isabelle
;  APPLICANT: Ellerman, Karen
;  APPLICANT: Stone, David J.
;  APPLICANT: Grosse, William M.
;  APPLICANT: Lepley, Denise M.
;  APPLICANT: Rieger, Daniel K.
;  APPLICANT: Burgess, Cathereine E.
;  APPLICANT: Casman, Stacie, J.
;  APPLICANT: Spytek, Kimberly A.
;  APPLICANT: Boldog, Ferenc L.
;  APPLICANT: Li, Li
;  APPLICANT: Padigaru, Muralidhara
;  APPLICANT: Mishra, Vishnu
;  APPLICANT: Shenoy, Suresh G.
;  APPLICANT: Rastelli, Luca

```
;  APPLICANT:  Tchernev, Velizar T.
;  APPLICANT:  Vernet, Corine A.M.
;  APPLICANT:  Zerhusen, Bryan D.
;  APPLICANT:  Malyankar, Uriel M.
;  APPLICANT:  Guo, Xiaojia
;  APPLICANT:  Miller, Charles E.
;  APPLICANT:  Gangolli, Esha A.
;  TITLE OF INVENTION: PROTEINS AND NUCLEIC ACIDS ENCODING SAME
;  FILE REFERENCE: 21402-214 CIP
;  CURRENT APPLICATION NUMBER: US/10/087,684
;  CURRENT FILING DATE:  2003-03-10
;  PRIOR APPLICATION NUMBER: 60/253,834
;  PRIOR FILING DATE: 2000-11-29
;  PRIOR APPLICATION NUMBER: 60/250,926
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: 60/264,180
;  PRIOR FILING DATE: 2001-01-25
;  PRIOR APPLICATION NUMBER: 60/274,194
;  PRIOR FILING DATE: 2001-03-08
;  PRIOR APPLICATION NUMBER: 60/313,656
;  PRIOR FILING DATE: 2001-08-20
;  PRIOR APPLICATION NUMBER: 60/327,456
;  PRIOR FILING DATE: 2001-10-05
;  NUMBER OF SEQ ID NOS: 220
;  SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 3
;   LENGTH: 2860
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (59)..(2857)
US-10-087-684-3

Alignment Scores:
Pred. No.:               1.72e-251    Length:         2860
Score:                   2484.00      Matches:        493
Percent Similarity:      68.30%       Conservative:   149
Best Local Similarity:   52.45%       Mismatches:     250
Query Match:             51.89%       Indels:         48
DB:                      13           Gaps:           18

US-09-970-944-2 (1-899) x US-10-087-684-3 (1-2860)

Qy        1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrp--- 19
Db       59 ATGGGGGCCCGGAGCGGAGCTCGGGGCGCGCTGCTGCTGGCACTGCTGCTCTGCTGGGAC 118

Qy       20 ------LeuArgGlySerGlyAlaGlnGln---SerAlaThrValAlaAsnProValPro 36
Db      119 CCGAGGCTGAGCCAAGCAGGCACTGATTCTGGCAGCGAGGTGCTCCCTGACTCCTTCCCG 178

Qy       37 GlyAlaAsnProAspLeuLeuProHisPheLeuValGluProGluAspValTyrIleVal 56
Db      179 TCAGCGCCAGCAGAGCCGCTGCCCTACTTCCTGCAGGAGCCACAGGACGCCTACATTGTG 238

Qy       57 LysAsnLysProValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLys 76
Db      239 AAGAACAAGCCTGTGGAGCTTCGCTGCCGCGCCTTCCCCGCCACACAGATCTACTTCAAG 298

Qy       77 CysAsnGlyGluTrpValArgGlnValAspHisValIleGluArgSerThrAspGlySer 96
Db      299 TGCAACGGCGAGTGGGTCAGCCAGAACGACCACGTCACACAGGAAGGCCTGGATGAGGCC 358

Qy       97 SerGlyGluProThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysVal 116
Db      359 ACCGGCCTGCGGGTGCGCGAGGTGCAGATCGAGGTGTCGCGGCAGCAGGTGGAGGAGCTC 418

Qy      117 PheGlyLeuGluGluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLys 136
Db      419 TTTGGGCTGGAGGATTACTGGTGCCAGTGCGTGGCCTGGAGCTCCGCAGGCACCACCAAG 478

Qy      137 SerGlnLysAlaTyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeu 156
Db      479 AGTCGCCGAGCCTACGTCCGCATCGCCTACCTGCGCAAGAACTTCGATCAGGAGCCTCTG 538

Qy      157 AlaLysGluValSerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIle 176
Db      539 GGCAAGGAGGTGCCCCTGGACCATGAGGTTCTCCTGCAGTGCCGCCCGCCGGAGGGGGTG 598

Qy      177 ProProAlaGluValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspPro 196
Db      599 CCTGTGGCCGAGGTGGAATGGCTCAAGAATGAGGATGTCATCGACCCCACCCAGGACACC 658

Qy      197 AsnValTyrIleThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThr 216
Db      659 AACTTCCTGCTCACCATCGACCACAACCTCATCATCCGCCAGGCCCGCCTGTCGGACACT 718

Qy      217 AlaAsnTyrThrCysValAlaLysAsnIleValAlaArgArgArgSerAlaSerAlaAla 236
Db      719 GCCAACTATACCTGCGTGGCCAAGAACATCGTGGCCAAACGCCGGAGCACCACTGCCACC 778

Qy      237 ValIleValTyrValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAla 256
Db      779 GTCATCGTCTACGTGAATGGCGGCTGGTCCAGCTGGGCAGAGTGGTCACCCTGCTCCAAC 838

Qy      257 SerCysGlyArgGlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsn 276
Db      839 CGCTGTGGCCGAGGCTGGCAGAAGCGCACCCGGACCTGCACCAACCCCGCTCCACTCAAC 898

Qy      277 GlyGlyAlaPheCysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuVal 296
Db      899 GGAGGGGCCTTCTGCGAGGGCCAG---GCATTCCAGAAGACCGCCTGCACCACCATCTGC 955

Qy      297 SerValAspGlySerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThr 316
Db      956 CCAGTCGATGGGGCGTGGACGGAGTGGAGCAAGTGGTCAGCCTGCAGCACTGAGTGTGCC 1015

Qy      317 HisTrpArgSerArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGln 336
Db     1016 CACTGGCGTAGCCGCGAGTGCATGGCGCCCCCACCCCAGAACGGAGGCCGTGACTGCAGC 1075

Qy      337 GlyThrAspLeuAspThrArgAsnCysThrSerAspLeuCysValHis---SerAlaSer 355
Db     1076 GGGACGCTGCTCGACTCTAAGAACTGCACAGATGGGCTGTGCATGCAACTGGAGGCCTCA 1135

Qy      356 GlyProGluAspValAlaLeuTyrValGlyLeu---IleAlaValAlaValCysLeuVal 374
Db     1136 GGG------GATGCGGCGCTGTATGCGGGGCTCGTGGTGGCCATCTTCGTGGTCGTGGCA 1189

Qy      375 LeuLeuLeuLeuValLeuIleLeuValTyrCysArgLysLysGluGlyLeuAspSerAsp 394
Db     1190 ATCCTCATGGCGGTGGGGGTGGTGGTGTACCGCCGCAACTGCCGTGACTTCGACACAGAC 1249

Qy      395 ValAlaAspSerSer---IleLeuThrSerGlyPheGlnProValSerIleLysProSer 413
Db     1250 ATCACTGACTCATCTGCTGCCCTGACTGGTGGTTTCCACCCCGTCAACTTTAAGACGGCA 1309

Qy      414 LysAlaAspAsnProHisLeuLeu------ThrIleGlnProAspLeuSerThrThrThr 431
Db     1310 AGGCCCAGTAACCCGCAGCTCCTACACCCCTCTGTGCCTCCTGACCTGACAGCCAGCGCC 1369

Qy      432 ---ThrTyrGlnGlySerLeuCysProArgGlnAspGlyProSerProLysPheGlnLeu 450
Db     1370 GGCATCTACCGCGGACCCGTGTATGCCCTGCAGGACTCC---ACCGACAAAATCCCCATG 1426

Qy      451 ThrAsnGlyHisLeuLeuSerProLeuGlyGlyGlyArgHisThrLeuHisHisSerSer 470
Db     1427 ACCAACTCTCCTCTGCTGGACCCCTTACCCAGCCTTAAGGTCAAGGTCTACAGCTCCAGC 1486

Qy      471 ProThr------------------SerGluAlaGluGluPheValSerArgLeuSerThr 484
Db     1487 ACCACGGGCTCTGGGCCAGGCCTGGCAGATGGGGCTGACCTGCTGGGGGTCTTGCCGCCT 1546

Qy      485 GlnAsnTyr-----------------------------------PheArgSer------ 490
```

```
Db       1547 GGCACATACCCTAGCGATTTCGCCCGGGACACCCACTTCCTGCACCTGCGCAGCGCCAGC 1606

Qy        491 -----------------------LeuProArgGlyThrSerAsnMetThrTyrGlyThr 502
                                     ||||||||          :::        ||||||
Db       1607 CTCGGTTCCCAGCAGCTCTTGGGCCTGCCCCGAGACCCAGGGAGCAGCGTCAGCGGCACC 1666

Qy        503 PheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeuLeuIlePro 522
              |||     ||||||||||||||    ||||||    |||||:::|||||||||:::|||
Db       1667 TTTGGCTGCCTGGGTGGAGGCTCAGCATCCCCGGCACAGGGGTCAGCTTGCTGGTGCCC 1726

Qy        523 ProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHisLysProGlu 542
                 ||||||||:::||||||    ||||||:::||||||    ::::::|||   |||
Db       1727 AATGGAGCCATTCCCCAGGGCAAGTTCTACGAGATGTATCTACTCATCAACAAGGCAGAA 1786

Qy        543 AspValArgLeuProLeuAla---GlyCysGlnThrLeuLeuSerProIleValSerCys 561
                       |||||||||:::    |||   ||||||:::|||||||||   |||:::|||
Db       1787 AGTACC---CTGCCGCTTTCAGAAGGGACCCAGACAGTATTGAGCCCCTCGGTGACCTGT 1843

Qy        562 GlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGlyGlu 581
              ||||||    |||:::||||||    ||||||||||||||    |||    ||||||   |||
Db       1844 GGACCCACAGGCCTCCTGCTGTGCCGCCCCGTCATCCTCACCATGCCCCACTGTGCCGAA 1903

Qy        582 ProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySerTrpGlu 601
                 |||         |||        :::||||||    |||:::    :::|||   ||||||
Db       1904 GTCAGTGCCCGTGACTGGATCTTTCAGCTCAAGACCCAGGCCCACCAGGGCCACTGGGAG 1963

Qy        602 GlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGlnLeuGlu 621
                 :::|||:::   |||   ||||||        :::         ||||||||||||||
Db       1964 ---GAGGTGGTGACCCTGGATGAGGAGACCCTGAACACACCCTGCTACTGCCAGCTGGAG 2020

Qy        622 AlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGluAla 641
                    ||||||::::::         :::|||||||||   :::         ||||||:::
Db       2021 CCCAGGGCCTGTCACATCCTGCTGGACCAGCTGGGCACCTACGTGTTCACGGGCGAGTCC 2080

Qy        642 LeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAlaCysThr 661
                 |||    :::|||   |||||||||:::|||     :::|||||||||      ||||||
Db       2081 TATTCCCGCTCAGCAGTCAAGCGGCTCCAGCTGGCCGTCTTCGCCCCCGCCCTCTGCACC 2140

Qy        662 SerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuLysGlu 681
              ||||||||||||::::::|||||||||||||||    ||||||      ||||||||||||
Db       2141 TCCCTGGAGTACAGCCTCCGGGTCTACTGCCTGGAGGACACGCCTGTAGCACTGAAGGAG 2200

Qy        682 ValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgValLeuHis 701
              |||::::::||||||:::    |||||||||    |||::::::||||||:::     |||
Db       2201 GTGCTGGAGCTGGAGCGGACTCTGGGCGGATACTTGGTGGAGGAGCCGAAACCGCTAATG 2260

Qy        702 PheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSerLeuTrp 721
              ||||||||||||||||||||||||||||||||||:::||||||:::|||    :::    |||
Db       2261 TTCAAGGACAGTTACCACAACCTGCGCCTCTCCCTCCATGACCTCCCCCATGCCCATTGG 2320

Qy        722 LysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsnGlyThr 741
              :::||||||||||||    ||||||||||||||||||||||||||||||:::|||:::
Db       2321 AGGAGCAAGCTGCTGGCCAAATACCAGGAGATCCCCTTCTATCACATTTGGAGTGGCAGC 2380

Qy        742 GlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSerAspLeu 761
              |||:::    ||||||||||||||||||||||||    |||    ::::::::::::|||
Db       2381 CAGAAGGCCCTCCACTGCACTTTCACCCTGGAGAGGCACAGCTTGGCCTCCACAGAGCTC 2440

Qy        762 AlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIleAsnPhe 781
                 ||||||:::   |||   |||||||||||||:::||||||   |||   ::::::
Db       2441 ACCTGCAAGATCTGCGTGCGGCAAGTGGAAGGGGAGGGCCAGATATTCCAGCTGCATACC 2500

Qy        782 AsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAlaGly--- 800
                 :::       :::|||      |||      |||      |||      |||
Db       2501 ACTCTGGCA---GAGACACCTGCTGGCTCCCTGGACACTCTCTGCTCTGCCCCTGGCAGC 2557

Qy        801 ---ValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
                 |||         :::||||||   |||||||||||||||      ||||||||||||
Db       2558 ACTGTCACCACCCAGCTGGGACCTTATGCCTTCAAGATCCCACTGTCCATCCGCCAGAAG 2617

Qy        820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839
              |||   :::|||||||||   |||      ||||||   |||||||||   |||||||||
Db       2618 ATATGCAACAGCCTAGATGCCCCCAACTCACGGGGCAATGACTGGCGGATGTTAGCACAG 2677

Qy        840 LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859
              ||||||   :::|||    :::||||:::::|||||||:::|||   |||||||||   :::
Db       2678 AAGCTCTCTATGGACCGGTACCTGAATTACTTTGCCACCAAAGCGAGCCCCACGGGTGTG 2737

Qy        860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879
              ||||||:::||||||||||||                :::|||:::|||:::    ||||||:::
Db       2738 ATCCTGGACCTCTGGGAAGCTCTGCAGCAGGACGATGGGGACCTCAACAGCCTGGCGAGT 2797

Qy        880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
              |||:::      :::|||:::   :::      :::      ::::::   :::|||
Db       2798 GCCTTGGAGGAGATGGGCAAGAGTGAGATGCTGGTGGCTGTGGCCACCGACGGGGACTGC 2857
```

RESULT 14
US-10-218-779-1
; Sequence 1, Application US/10218779
; Publication No. US20040029222A1
; GENERAL INFORMATION:
;  APPLICANT: Edinger, Shlomit
;  APPLICANT: MacDougall, John
;  APPLICANT: Millet, Isabelle
;  APPLICANT: Ellerman, Karen
;  APPLICANT: Stone, David
;  APPLICANT: Gerlach, Valerie
;  APPLICANT: Grosse, William
;  APPLICANT: Alsobrook II, John
;  APPLICANT: Lepley, Denise
;  APPLICANT: Rieger, Daniel
;  APPLICANT: Burgess, Catherine
;  APPLICANT: Casman, Stacie
;  APPLICANT: Spytek, Kimberly
;  APPLICANT: Boldog, Ferenc
;  APPLICANT: Li, Li
;  APPLICANT: Padigaru, Muralidhara
;  APPLICANT: Mishra, Vishnu
;  APPLICANT: Patturajan, Meera
;  APPLICANT: Shenoy, Suresh
;  APPLICANT: Rastelli, Luca
;  APPLICANT: Tchernev, Velizar
;  APPLICANT: Vernet, Corine
;  APPLICANT: Zerhusen, Bryan
;  APPLICANT: Malyankar, Uriel
;  APPLICANT: Guo, Xiaojia
;  APPLICANT: Miller, Charles
;  APPLICANT: Gangolli, Esha
;  TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
;  FILE REFERENCE: 21402-214
;  CURRENT APPLICATION NUMBER: US/10/218,779
;  CURRENT FILING DATE: 2002-08-14
;  PRIOR APPLICATION NUMBER: 60/253,834
;  PRIOR FILING DATE: 2000-11-29
;  PRIOR APPLICATION NUMBER: 60/250,-926
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: 60/264,180
;  PRIOR FILING DATE: 2001-01-25
;  PRIOR APPLICATION NUMBER: 60/313,656
;  PRIOR FILING DATE: 2001-08-20
;  PRIOR APPLICATION NUMBER: 60/327,456
;  PRIOR FILING DATE: 2001-10-05
;  NUMBER OF SEQ ID NOS: 216
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 2860
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-218-779-1

Alignment Scores:

| | | | |
|---|---|---|---|
| Pred. No.: | 1.72e-251 | Length: | 2860 |
| Score: | 2484.00 | Matches: | 493 |
| Percent Similarity: | 68.30% | Conservative: | 149 |
| Best Local Similarity: | 52.45% | Mismatches: | 250 |
| Query Match: | 51.89% | Indels: | 48 |

```
DB:                13          Gaps:          18
US-09-970-944-2 (1-899) x US-10-218-779-1 (1-2860)

Qy      1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrp--- 19
Db     59 ATGGGGGCCCGGAGCGGAGCTCGGGGCGCGCTGCTGCTGGCACTGCTGCTCTGCTGGGAC 118

Qy     20 ------LeuArgGlySerGlyAlaGlnGln---SerAlaThrValAlaAsnProValPro 36
Db    119 CCGAGGCTGAGCCAAGCAGGCACTGATTCTGGCAGCGAGGTGCTCCCTGACTCCTTCCCG 178

Qy     37 GlyAlaAsnProAspLeuLeuProHisPheLeuValGluProGluAspValTyrIleVal 56
Db    179 TCAGCGCCAGCAGAGCCGCTGCCCTACTTCCTGCAGGAGCCACAGGACGCCTACATTGTG 238

Qy     57 LysAsnLysProValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLys 76
Db    239 AAGAACAAGCCTGTGGAGCTCCGCTGCCGCGCCTTCCCCGCCACACAGATCTACTTCAAG 298

Qy     77 CysAsnGlyGluTrpValArgGlnValAspHisValIleGluArgSerThrAspGlySer 96
Db    299 TGCAACGGCGAGTGGGTCAGCCAGAACGACCACGTCACACAGGAAGGCCTGGATGAGGCC 358

Qy     97 SerGlyGluProThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysVal 116
Db    359 ACCGGTCTGCGGGTGCGCGAGGTGCAGATCGAGGTGTCGCGGCAGCAGGTGGAGGAGCTC 418

Qy    117 PheGlyLeuGluGluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLys 136
Db    419 TTTGGGCTGGAGGATTACTGGTGCCAGTGCGTGGCCTGGAGCTCCGCGGGCACCACCAAG 478

Qy    137 SerGlnLysAlaTyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeu 156
Db    479 AGTCGCCGAGCCTACGTCCGCATCGCCTACCTGCGCAAGAACTTCGATCAGGAGCCTCTG 538

Qy    157 AlaLysGluValSerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIle 176
Db    539 GGCAAGGAGGTGCCCCTGGACCATGAGGTTCTCCTGCAGTGCCGCCCGCCGGAGGGGGTG 598

Qy    177 ProProAlaGluValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspPro 196
Db    599 CCTGTGGCCGAGGTGGAATGGCTCAAGAATGAGGATGTCATCGACCCCACCCAGGACACC 658

Qy    197 AsnValTyrIleThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThr 216
Db    659 AACTTCCTGCTCACCATCGACCACAACCTCATCATCCGCCAGGCCCGCCTGTCGGACACT 718

Qy    217 AlaAsnTyrThrCysValAlaLysAsnIleValAlaArgArgArgSerAlaSerAlaAla 236
Db    719 GCCAACTATACCTGCGTGGCCAAGAACATCGTGGCCAAACGCCGGAGCACCACTGCCACC 778

Qy    237 ValIleValTyrValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAla 256
Db    779 GTCATCGTCTACGTGAATGGCGGCTGGTCCAGCTGGGCAGAGTGGTCACCCTGCTCCAAC 838

Qy    257 SerCysGlyArgGlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsn 276
Db    839 CGCTGTGGCCGAGGCTGGCAGAAGCGCACCCGGACCTGCACCAACCCCGCTCCACTCAAC 898

Qy    277 GlyGlyAlaPheCysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuVal 296
Db    899 GGAGGGGCCTTCTGCGAGGGCCAG---GCATTCCAGAAGACCGCCTGCACCACCATCTGC 955

Qy    297 SerValAspGlySerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThr 316
Db    956 CCAGTCGATGGGGCGTGGACGGAGTGGAGCAAGTGGTCAGCCTGCAGCACTGAGTGTGCC 1015

Qy    317 HisTrpArgSerArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGln 336
Db   1016 CACTGGCGTAGCCGCGAGTGCATGGCGCCCCCACCCCAGAACGGAGGCCGTGACTGCAGC 1075

Qy    337 GlyThrAspLeuAspThrArgAsnCysThrSerAspLeuCysValHis---SerAlaSer 355
Db   1076 GGGACGCTGCTCGACTCTAAGAACTGCACAGATGGGCTGTGCATGCAACTGGAGGCCTCA 1135

Qy    356 GlyProGluAspValAlaLeuTyrValGlyLeu---IleAlaValAlaValCysLeuVal 374
Db   1136 GGG------GATGCGGCGCTGTATGCGGGGCTCGTGGTGGCCATCTTCGTGGTCGTGGCA 1189

Qy    375 LeuLeuLeuLeuValLeuIleLeuValTyrCysArgLysLysGluGlyLeuAspSerAsp 394
Db   1190 ATCCTCATGGCGGTGGGGGTGGTGGTGTACCGCCGCAACTGCCGTGACTTCGACACAGAC 1249

Qy    395 ValAlaAspSerSer---IleLeuThrSerGlyPheGlnProValSerIleLysProSer 413
Db   1250 ATCACTGACTCATCTGCTGCCCTGACTGGTGGTTTCCACCCCGTCAACTTTAAGACGGCA 1309

Qy    414 LysAlaAspAsnProHisLeuLeu------ThrIleGlnProAspLeuSerThrThrThr 431
Db   1310 AGGCCCAGTAACCCGCAGCTCCTACACCCCTCTGTGCCTCCTGACCTGACAGCCAGCGCC 1369

Qy    432 ---ThrTyrGlnGlySerLeuCysProArgGlnAspGlyProSerProLysPheGlnLeu 450
Db   1370 GGCATCTACCGCGGACCCGTGTATGCCCTGCAGGACTCC---ACCGACAAAATCCCCATG 1426

Qy    451 ThrAsnGlyHisLeuLeuSerProLeuGlyGlyGlyArgHisThrLeuHisHisSerSer 470
Db   1427 ACCAACTCTCCTCTGCTGGACCCCTTACCCAGCCTTAAGGTCAAGGTCTACAGCTCCAGC 1486

Qy    471 ProThr-----------------SerGluAlaGluGluPheValSerArgLeuSerThr 484
Db   1487 ACCACGGGCTCTGGGCCAGGCCTGGCAGATGGGGCTGACCTGCTGGGGGTCTTGCCGCCT 1546

Qy    485 GlnAsnTyr----------------------------------PheArgSer------ 490
Db   1547 GGCACATACCCTAGCGATTTCGCCCGGGACACCCACTTCCTGCACCTGCGCAGCGCCAGC 1606

Qy    491 -----------------------LeuProArgGlyThrSerAsnMetThrTyrGlyThr 502
Db   1607 CTCGGTTCCCAGCAGCTCTTGGGCCTGCCCCGAGACCCAGGGAGCAGCGTCAGCGGCACC 1666

Qy    503 PheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeuLeuIlePro 522
Db   1667 TTTGGCTGCCTGGGTGGGAGGCTCAGCATCCCCGGCACAGGGGTCAGCTTGCTGGTGCCC 1726

Qy    523 ProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHisLysProGlu 542
Db   1727 AATGGAGCCATTCCCCAGGGCAAGTTCTACGAGATGTATCTACTCATCAACAAGGCAGAA 1786

Qy    543 AspValArgLeuProLeuAla---GlyCysGlnThrLeuLeuSerProIleValSerCys 561
Db   1787 AGTACC---CTGCCGCTTTCAGAAGGGACCCAGACAGTATTGAGCCCCTCGGTGACCTGT 1843

Qy    562 GlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGlyGlu 581
Db   1844 GGACCCACAGGCCTCCTGCTGTGCCGCCCCGTCATCCTCACCATGCCCCACTGTGCCGAA 1903

Qy    582 ProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySerTrpGlu 601
Db   1904 GTCAGTGCCCGTGACTGGATCTTTCAGCTCAAGACCCAGGCCCACCAGGGCCACTGGGAG 1963

Qy    602 GlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGlnLeuGlu 621
Db   1964 ---GAGGTGGTGACCCTGGATGAGGAGACCCTGAACACACCCTGCTACTGCCAGCTGGAG 2020

Qy    622 AlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGluAla 641
Db   2021 CCCAGGGCCTGTCACATCCTGCTGGACCAGCTGGGCACCTACGTGTTCACGGGCGAGTCC 2080

Qy    642 LeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAlaCysThr 661
Db   2081 TATTCCCGCTCAGCAGTCAAGCGGCTCCAGCTGGCCGTCTTCGCCCCCGCCCTCTGCACC 2140

Qy    662 SerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuLysGlu 681
```

```
Db       2141 TCCCTGGAGTACAGCCTCCGGGTCTACTGCCTGGAGGACACGCCTGTAGCACTGAAGGAG 2200

Qy        682 ValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgValLeuHis 701
Db       2201 GTGCTGGAGCTGGAGCGGACTCTGGGCGGATACTTGGTGGAGGAGCCGAAACCGCTAATG 2260

Qy        702 PheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSerLeuTrp 721
Db       2261 TTCAAGGACAGTTACCACAACCTGCGCCTCTCCCTCCATGACCTCCCCCATGCCCATTGG 2320

Qy        722 LysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsnGlyThr 741
Db       2321 AGGAGCAAGCTGCTGGCCAAATACCAGGAGATCCCCTTCTATCACATTTGGAGTGGCAGC 2380

Qy        742 GlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSerAspLeu 761
Db       2381 CAGAAGGCCCTCCACTGCACTTTCACCCTGGAGAGGCACAGCTTGGCCTCCACAGAGCTC 2440

Qy        762 AlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIleAsnPhe 781
Db       2441 ACCTGCAAGATCTGCGTGCGGCAAGTGGAAGGGGAGGGCCAGATATTCCAGCTGCATACC 2500

Qy        782 AsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAlaGly--- 800
Db       2501 ACTCTGGCA---GAGACACCTGCTGGCTCCCTGGACACTCTCTGCTCTGCCCCTGGCAGC 2557

Qy        801 ---ValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
Db       2558 ACTGTCACCACCCAGCTGGGACCTTATGCCTTCAAGATCCCACTGTCCATCCGCCAGAAG 2617

Qy        820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839
Db       2618 ATATGCAACAGCCTAGATGCCCCCAACTCACGGGGCAATGACTGGCGGATGTTAGCACAG 2677

Qy        840 LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859
Db       2678 AAGCTCTCTATGGACCGGTACCTGAATTACTTTGCCACCAAAGCGAGCCCCACGGGTGTG 2737

Qy        860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879
Db       2738 ATCCTGGACCTCTGGGAAGCTCTGCAGCAGGACGATGGGGACCTCAACAGCCTGGCGAGT 2797

Qy        880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
Db       2798 GCCTTGGAGGAGATGGGCAAGAGTGAGATGCTGGTGGCTGTGGCCACCGACGGGGACTGC 2857

RESULT 15
US-10-218-779-3
; Sequence 3, Application US/10218779
; Publication No. US20040029222A1
; GENERAL INFORMATION:
;  APPLICANT: Edinger, Shlomit
;  APPLICANT:  MacDougall, John
;  APPLICANT:  Millet, Isabelle
;  APPLICANT:  Ellerman, Karen
;  APPLICANT:  Stone, David
;  APPLICANT:  Gerlach, Valerie
;  APPLICANT:  Grosse, William
;  APPLICANT:  Alsobrook II, John
;  APPLICANT:  Lepley, Denise
;  APPLICANT:  Rieger, Daniel
;  APPLICANT:  Burgess, Catherine
;  APPLICANT:  Casman, Stacie
;  APPLICANT:  Spytek, Kimberly
;  APPLICANT:  Boldog, Ferenc
;  APPLICANT:  Li, Li
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Mishra, Vishnu
;  APPLICANT:  Patturajan, Meera
;  APPLICANT:  Shenoy, Suresh
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Tchernev, Velizar
;  APPLICANT:  Vernet, Corine
;  APPLICANT:  Zerhusen, Bryan
;  APPLICANT:  Malyankar, Uriel
;  APPLICANT:  Guo, Xiaojia
;  APPLICANT:  Miller, Charles
;  APPLICANT:  Gangolli, Esha
;  TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
;  FILE REFERENCE: 21402-214
;  CURRENT APPLICATION NUMBER: US/10/218,779
;  CURRENT FILING DATE:  2002-08-14
;  PRIOR APPLICATION NUMBER: 60/253,834
;  PRIOR FILING DATE: 2000-11-29
;  PRIOR APPLICATION NUMBER: 60/250,-926
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: 60/264,180
;  PRIOR FILING DATE: 2001-01-25
;  PRIOR APPLICATION NUMBER: 60/313,656
;  PRIOR FILING DATE: 2001-08-20
;  PRIOR APPLICATION NUMBER: 60/327,456
;  PRIOR FILING DATE: 2001-10-05
;  NUMBER OF SEQ ID NOS: 216
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 2860
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-218-779-3

Alignment Scores:
Pred. No.:              1.72e-251     Length:          2860
Score:                  2484.00       Matches:         493
Percent Similarity:     68.30%        Conservative:    149
Best Local Similarity:  52.45%        Mismatches:      250
Query Match:            51.89%        Indels:          48
DB:                     13            Gaps:            18

US-09-970-944-2 (1-899) x US-10-218-779-3 (1-2860)

Qy          1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrp--- 19
Db         59 ATGGGGGCCCGGAGCGGAGCTCGGGGCGCGCTGCTGCTGGCACTGCTGCTCTGCTGGGAC 118

Qy         20 ------LeuArgGlySerGlyAlaGlnGln---SerAlaThrValAlaAsnProValPro 36
Db        119 CCGAGGCTGAGCCAAGCAGGCACTGATTCTGGCAGCGAGGTGCTCCCTGACTCCTTCCCG 178

Qy         37 GlyAlaAsnProAspLeuLeuProHisPheLeuValGluProGluAspValTyrIleVal 56
Db        179 TCAGCGCCAGCAGAGCCGCTGCCCTACTTCCTGCAGGAGCCACAGGACGCCTACATTGTG 238

Qy         57 LysAsnLysProValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLys 76
Db        239 AAGAACAAGCCTGTGGAGCTTCGCTGCCGCGCCTTCCCCGCCACACAGATCTACTTCAAG 298

Qy         77 CysAsnGlyGluTrpValArgGlnValAspHisValIleGluArgSerThrAspGlySer 96
Db        299 TGCAACGGCGAGTGGGTCAGCCAGAACGACCACGTCACACAGGAAGGCCTGGATGAGGCC 358

Qy         97 SerGlyGluProThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysVal 116
Db        359 ACCGGCCTGCGGGTGCGCGAGGTGCAGATCGAGGTGTCGCGGCAGCAGGTGGAGGAGCTC 418

Qy        117 PheGlyLeuGluGluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLys 136
Db        419 TTTGGGCTGGAGGATTACTGGTGCCAGTGCGTGGCCTGGAGCTCCGCAGGCACCACCAAG 478

Qy        137 SerGlnLysAlaTyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeu 156
Db        479 AGTCGCCGAGCCTACGTCCGCATCGCCTACCTGCGCAAGAACTTCGATCAGGAGCCTCTG 538

Qy        157 AlaLysGluValSerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIle 176
Db        539 GGCAAGGAGGTGCCCCTGGACCATGAGGTTCTCCTGCAGTGCCGCCCGCCGGAGGGGGTG 598
```

```
Qy    177 ProProAlaGluValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspPro 196
Db    599 CCTGTGGCCGAGGTGGAATGGCTCAAGAATGAGGATGTCATCGACCCCACCCAGGACACC 658
Qy    197 AsnValTyrIleThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThr 216
Db    659 AACTTCCTGCTCACCATCGACCACAACCTCATCATCCGCCAGGCCCGCCTGTCGGACACT 718
Qy    217 AlaAsnTyrThrCysValAlaLysAsnIleValAlaArgArgArgSerAlaSerAlaAla 236
Db    719 GCCAACTATACCTGCGTGGCCAAGAACATCGTGGCCAAACGCCGGAGCACCACTGCCACC 778
Qy    237 ValIleValTyrValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAla 256
Db    779 GTCATCGTCTACGTGAATGGCGGCTGGTCCAGCTGGGCAGAGTGGTCACCCTGCTCCAAC 838
Qy    257 SerCysGlyArgGlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsn 276
Db    839 CGCTGTGGCCGAGGCTGGCAGAAGCGCACCCGGACCTGCACCAACCCCGCTCCACTCAAC 898
Qy    277 GlyGlyAlaPheCysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuVal 296
Db    899 GGAGGGGCCTTCTGCGAGGGCCAG---GCATTCCAGAAGACCGCCTGCACCACCATCTGC 955
Qy    297 SerValAspGlySerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThr 316
Db    956 CCAGTCGATGGGGCGTGGACGGAGTGGAGCAAGTGGTCAGCCTGCAGCACTGAGTGTGCC 1015
Qy    317 HisTrpArgSerArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGln 336
Db   1016 CACTGGCGTAGCCGCGAGTGCATGGCGCCCCCACCCCAGAACGGAGGCCGTGACTGCAGC 1075
Qy    337 GlyThrAspLeuAspThrArgAsnCysThrSerAspLeuCysValHis---SerAlaSer 355
Db   1076 GGGACGCTGCTCGACTCTAAGAACTGCACAGATGGGCTGTGCATGCAACTGGAGGCCTCA 1135
Qy    356 GlyProGluAspValAlaLeuTyrValGlyLeu---IleAlaValAlaValCysLeuVal 374
Db   1136 GGG------GATGCGGCGCTGTATGCGGGGCTCGTGGTGGCCATCTTCGTGGTCGTGGCA 1189
Qy    375 LeuLeuLeuLeuValLeuIleLeuValTyrCysArgLysLysGluGlyLeuAspSerAsp 394
Db   1190 ATCCTCATGGCGGTGGGGGTGGTGGTGTACCGCCGCAACTGCCGTGACTTCGACACAGAC 1249
Qy    395 ValAlaAspSerSer---IleLeuThrSerGlyPheGlnProValSerIleLysProSer 413
Db   1250 ATCACTGACTCATCTGCTGCCCTGACTGGTGGTTTCCACCCCGTCAACTTTAAGACGGCA 1309
Qy    414 LysAlaAspAsnProHisLeuLeu------ThrIleGlnProAspLeuSerThrThrThr 431
Db   1310 AGGCCCAGTAACCCGCAGCTCCTACACCCCTCTGTGCCTCCTGACCTGACAGCCAGCGCC 1369
Qy    432 ---ThrTyrGlnGlySerLeuCysProArgGlnAspGlyProSerProLysPheGlnLeu 450
Db   1370 GGCATCTACCGCGGACCCGTGTATGCCCTGCAGGACTCC---ACCGACAAAATCCCCATG 1426
Qy    451 ThrAsnGlyHisLeuLeuSerProLeuGlyGlyGlyArgHisThrLeuHisHisSerSer 470
Db   1427 ACCAACTCTCCTCTGCTGGACCCCTTACCCAGCCTTAAGGTCAAGGTCTACAGCTCCAGC 1486
Qy    471 ProThr------------------SerGluAlaGluGluPheValSerArgLeuSerThr 484
Db   1487 ACCACGGGCTCTGGGCCAGGCCTGGCAGATGGGGCTGACCTGCTGGGGGTCTTGCCGCCT 1546
Qy    485 GlnAsnTyr-----------------------------------PheArgSer------ 490
Db   1547 GGCACATACCCTAGCGATTTCGCCCGGGACACCCACTTCCTGCACCTGCGCAGCGCCAGC 1606
Qy    491 ------------------------LeuProArgGlyThrSerAsnMetThrTyrGlyThr 502
Db   1607 CTCGGTTCCCAGCAGCTCTTGGGCCTGCCCCGAGACCCAGGGAGCAGCGTCAGCGGCACC 1666
Qy    503 PheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeuLeuIlePro 522
Db   1667 TTTGGCTGCCTGGGTGGGAGGCTCAGCATCCCCGGCACAGGGGTCAGCTTGCTGGTGCCC 1726
Qy    523 ProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHisLysProGlu 542
Db   1727 AATGGAGCCATTCCCCAGGGCAAGTTCTACGAGATGTATCTACTCATCAACAAGGCAGAA 1786
Qy    543 AspValArgLeuProLeuAla---GlyCysGlnThrLeuLeuSerProIleValSerCys 561
Db   1787 AGTACC---CTGCCGCTTTCAGAAGGGACCCAGACAGTATTGAGCCCCTCGGTGACCTGT 1843
Qy    562 GlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGlyGlu 581
Db   1844 GGACCCACAGGCCTCCTGCTGTGCCGCCCCGTCATCCTCACCATGCCCCACTGTGCCGAA 1903
Qy    582 ProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySerTrpGlu 601
Db   1904 GTCAGTGCCCGTGACTGGATCTTTCAGCTCAAGACCCAGGCCCACCAGGGCCACTGGGAG 1963
Qy    602 GlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGlnLeuGlu 621
Db   1964 ---GAGGTGGTGACCCTGGATGAGGAGACCCTGAACACACCCTGCTACTGCCAGCTGGAG 2020
Qy    622 AlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGluAla 641
Db   2021 CCCAGGGCCTGTCACATCCTGCTGGACCAGCTGGGCACCTACGTGTTCACGGGCGAGTCC 2080
Qy    642 LeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAlaCysThr 661
Db   2081 TATTCCCGCTCAGCAGTCAAGCGGCTCCAGCTGGCCGTCTTCGCCCCCGCCCTCTGCACC 2140
Qy    662 SerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuLysGlu 681
Db   2141 TCCCTGGAGTACAGCCTCCGGGTCTACTGCCTGGAGGACACGCCTGTAGCACTGAAGGAG 2200
Qy    682 ValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgValLeuHis 701
Db   2201 GTGCTGGAGCTGGAGCGGACTCTGGGCGGATACTTGGTGGAGGAGCCGAAACCGCTAATG 2260
Qy    702 PheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSerLeuTrp 721
Db   2261 TTCAAGGACAGTTACCACAACCTGCGCCTCTCCCTCCATGACCTCCCCCATGCCCATTGG 2320
Qy    722 LysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsnGlyThr 741
Db   2321 AGGAGCAAGCTGCTGGCCAAATACCAGGAGATCCCCTTCTATCACATTTGGAGTGGCAGC 2380
Qy    742 GlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSerAspLeu 761
Db   2381 CAGAAGGCCCTCCACTGCACTTTCACCCTGGAGAGGCACAGCTTGGCCTCCACAGAGCTC 2440
Qy    762 AlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIleAsnPhe 781
Db   2441 ACCTGCAAGATCTGCGTGCGGCAAGTGGAAGGGGAGGGCCAGATATTCCAGCTGCATACC 2500
Qy    782 AsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAlaGly--- 800
Db   2501 ACTCTGGCA---GAGACACCTGCTGGCTCCCTGGACACTCTCTGCTCTGCCCCTGGCAGC 2557
Qy    801 ---ValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
Db   2558 ACTGTCACCACCCAGCTGGGACCTTATGCCTTCAAGATCCCACTGTCCATCCGCCAGAAG 2617
Qy    820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839
Db   2618 ATATGCAACAGCCTAGATGCCCCCAACTCACGGGGCAATGACTGGCGGATGTTAGCACAG 2677
Qy    840 LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859
Db   2678 AAGCTCTCTATGGACCGGTACCTGAATTACTTTGCCACCAAAGCGAGCCCCACGGGTGTG 2737
Qy    860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879
```

```
Db       2738 ATCCTGGACCTCTGGGAAGCTCTGCAGCAGGACGATGGGGACCTCAACAGCCTGGCGAGT 2797

Qy        880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
              |||:::      :::|||:::   :::      :::         ::::::   :::|||
Db       2798 GCCTTGGAGGAGATGGGCAAGAGTGAGATGCTGGTGGCTGTGGCCACCGACGGGGACTGC 2857
```

Search completed: October  6, 2004, 22:57:28
Job time : 1147 secs

OM protein - protein search, using sw model

```
Run on:          October  6, 2004, 17:53:42 ; Search time 29 Seconds
                                              (without alignments)
                                              2981.937 Million cell updates/sec

Title:           US-09-970-944-2
Perfect score:   4787
Sequence:        1 MAVRPGLWPALLGIVLAAWL..........AVAGLGQPDAGLFTVSEAEC 899

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters:      283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :        PIR_78:*
                 1:  pir1:*
                 2:  pir2:*
                 3:  pir3:*
                 4:  pir4:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 950.5 | 19.9 | 919 | 2 | T32541 | unc-5 protein - Ca |
| 2 | 950.5 | 19.9 | 947 | 1 | B44294 | unc-5 protein, lon |
| 3 | 280 | 5.8 | 1584 | 2 | T00026 | brain-specific ang |
| 4 | 273.5 | 5.7 | 1172 | 1 | TSHUP2 | thrombospondin 2 p |
| 5 | 273.5 | 5.7 | 1172 | 2 | A42587 | thrombospondin 2 p |
| 6 | 270.5 | 5.7 | 1074 | 2 | JC5928 | semaphorin F precu |
| 7 | 267.5 | 5.6 | 984 | 2 | T00326 | hypothetical prote |
| 8 | 267.5 | 5.6 | 1522 | 2 | T00028 | brain-specific ang |
| 9 | 263.5 | 5.5 | 1444 | 2 | T18856 | angiogenesis inhib |
| 10 | 258 | 5.4 | 1572 | 2 | T00027 | brain-specific ang |
| 11 | 248 | 5.2 | 1170 | 2 | A40558 | thrombospondin 1 p |
| 12 | 246 | 5.1 | 1170 | 1 | TSHUP1 | thrombospondin 1 p |
| 13 | 243.5 | 5.1 | 1178 | 1 | A39804 | thrombospondin pre |
| 14 | 220.5 | 4.6 | 1651 | 2 | T14160 | transmembrane rece |
| 15 | 217.5 | 4.5 | 1612 | 2 | T30805 | duttl protein - mo |
| 16 | 211.5 | 4.4 | 1265 | 1 | A37967 | neural cell adhesi |
| 17 | 208.5 | 4.4 | 469 | 1 | S29126 | properdin precurso |
| 18 | 206.5 | 4.3 | 254 | 2 | T15952 | hypothetical prote |
| 19 | 202.5 | 4.2 | 437 | 2 | S05478 | properdin - mouse |
| 20 | 198.5 | 4.1 | 788 | 2 | T25061 | hypothetical prote |
| 21 | 188 | 3.9 | 423 | 2 | T29549 | hypothetical prote |
| 22 | 185 | 3.9 | 1344 | 2 | T14316 | rig-1 protein - mo |
| 23 | 184 | 3.8 | 957 | 2 | T15976 | hypothetical prote |
| 24 | 183.5 | 3.8 | 1273 | 2 | T42405 | sax-3 protein - Ca |
| 25 | 183.5 | 3.8 | 1863 | 2 | S46217 | protein-tyrosine-p |
| 26 | 176.5 | 3.7 | 860 | 2 | T16892 | hypothetical prote |
| 27 | 175.5 | 3.7 | 1736 | 2 | A47747 | tight junction pro |
| 28 | 170.5 | 3.6 | 1745 | 2 | A46431 | tight junction-ass |
| 29 | 169 | 3.5 | 934 | 1 | A34372 | complement C6 prec |
| 30 | 167.5 | 3.5 | 1907 | 2 | S50893 | protein-tyrosine-p |
| 31 | 166 | 3.5 | 837 | 2 | T00355 | hypothetical prote |
| 32 | 160.5 | 3.4 | 550 | 2 | T47158 | hypothetical prote |
| 33 | 152.5 | 3.2 | 1499 | 2 | I50212 | protein-tyrosine-p |
| 34 | 152 | 3.2 | 951 | 2 | T00017 | gene ADAMTS-1 prot |
| 35 | 152 | 3.2 | 1501 | 2 | I58148 | protein-tyrosine-p |
| 36 | 151 | 3.2 | 2165 | 2 | T21371 | hypothetical prote |
| 37 | 150 | 3.1 | 1277 | 2 | T30532 | neural cell adhesi |
| 38 | 149 | 3.1 | 654 | 2 | T29247 | hypothetical prote |
| 39 | 148.5 | 3.1 | 946 | 1 | A47299 | ror-related recept |
| 40 | 148 | 3.1 | 1091 | 2 | A58532 | glial cell membran |
| 41 | 147 | 3.1 | 843 | 1 | A27340 | complement C7 prec |
| 42 | 147 | 3.1 | 888 | 2 | S23065 | ufo protein - mous |
| 43 | 146.5 | 3.1 | 584 | 1 | C8HUA | complement C8 alph |
| 44 | 146 | 3.0 | 1051 | 2 | A39712 | kinase-like protei |
| 45 | 145.5 | 3.0 | 2029 | 1 | TDFFLK | protein-tyrosine-p |

ALIGNMENTS

```
RESULT 1
T32541
unc-5 protein - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 31-Jan-2000
C;Accession: T32541
R;Latreille, P.
submitted to the EMBL Data Library, December 1997
A;Description: The sequence of C. elegans cosmid B0273.
A;Reference number: Z21187
A;Accession: T32541
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-919 <LAT>
A;Cross-references: EMBL:AF036698; PIDN:AAB88355.1; GSPDB:GN00022; CESP:B0273.4a
A;Experimental source: strain Bristol N2; clone B0273
C;Genetics:
A;Gene: unc-5; CESP:B0273.4a
A;Map position: 4
A;Introns: 41/3; 108/1; 142/3; 201/1; 323/2; 553/1; 858/3
C;Superfamily: unc-5 protein; immunoglobulin homology; SH3 homology; thrombospondin type

  Query Match             19.9%;  Score 950.5;  DB 2;  Length 919;
  Best Local Similarity   28.5%;  Pred. No. 1.6e-61;
  Matches  265;  Conservative  165;  Mismatches  375;  Indels  125;  Gaps   31;

Qy         49 EPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQVDHVIER--STDGSSGEPTMEVRI 106

Db          9 QPKSGYVIRNKPLRLQCRANHATKIRYKCSSKWID--DSRIEKLIGTDSTSGVGYIDASV 66

Qy        107 NVSRQQVEKVFGLEEYWCQCVAWSSSG-----TTKSQKAYIRIARLRKNFEQEPLAKEVS 161

Db         67 DISRIDVDTSGHVDAFQCQCYA---SGDDDQDVVASDVATVHLAYMRKHFLKSPVAQRVQ 123

Qy        162 LEQGIVLPCRPPEGIPPAEVEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTC 221

Db        124 EGTTLQLPCQAPESDPKAELTWYKDGVVVQP--DANVIRASDGSLIMSAARLSDSGNYTC 181

Qy        222 VAKNIVARRRSASAAVIVYVNGGWSTWTEW-SVCSASCG--------------RGWQKR 265

Db        182 EATNVANSRKTDPVEVQIYVDGGWSEWSPWIGTCHVDCPLLRQHAHRIRDPHDVLPHQRR 241

Qy        266 SRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDGSWSPWSKWSACGLDCTHWRSRECSD 325

Db        242 TRTCNNPAPLNDGEYCKGEEEMTRS-CKVPCKLDGGWSSWSDWSACSSSCHRYRTRACTV 300

Qy        326 PAPRNGGEECQGTDLDTRNCTSDLCVHSASG--PEDVALYVGLIAVAVCLVLLLLVLILV 383

Db        301 PPPMNGGQPCFGDDLMTQECPAQLCTADSSRIVISDTAVYGSVASIFIVASFILAILAMF 360

Qy        384 YCR----------KKEGLDSDVADS-------------------------SILTSGFQP 407
```

```
Db        361 CCKRGNSKKSKPLKPQKMNSEKAGGIYYSEPPGVRRLLLEHQHGTLLGEKISSCSQYFEP 420

Qy        408 VSIKPSKADNPHLLTIQPDLSTTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLH 467

Db        421 PPL-------PHSTTLR---SGKSAFSGYSSTRNAGSRAALIQECSSSSSGSGGKRTMLR 470

Qy        468 HSSPT-SEAEEFVSRLSTQNYFRSLPRGTS-NMTYGTFNFLGGRLMIPNTGISLLIPPDA 525

Db        471 TSSSNCSDDDNYATLYDYMEDKSVLGLDTSQNIVAAQIDSNGARLSLSKSGARLIVPELA 530

Qy        526 IPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGV-------LLTRPVILAMDH 578

Db        531 VEGEKM--LYLAVSDTLTDQPHLKPIESALSPVIVIGQCDVSMSAHDNILRRPVVVSFRH 588

Qy        579 CGEPSP-DSWSLRLKKQSCEGSWEQDVLHLGEEAPSHLYYCQLEASA--------CYVFT 629

Db        589 CASTFPRDNWQFTL--YADEGSGWQKAVTIGEENLNTNMFVQFEQPGKKNDGFGWCHVMT 646

Qy        630 EQLGRFALVGEAL--SVAAAKRLKLLLFAPVACTSLE--YNIRVYCLHDTHDALKEVVQL 685

Db        647 YSLARLMLAGHPRRNSLSAAKRVHLAVFGPTEMSAYRRPFELRVYCVPETGAAMESVWKQ 706

Qy        686 EKQLGGQLIQEPR--VLHFKDSYHNLRLSIHDV-PSSLWKSKLLVSYQEIPFYHIWNGTQ 742

Db        707 ED--GSRLLCESNDFILNEKG---NLCICIEDVIPGFSCDGPEVVEISETQHRFV---AQ 758

Qy        743 RYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVP 802

Db        759 NGLHCSLKFRPKEINGSQFSTRVIVYQKASSTEPMVM--EVSNEPELYDATSEEREKGSV 816

Qy        803 ALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLDSHLSFFASKP--SPTAMI 860

Db        817 CV----EFRLPFGVKDELARLLDMPNESHSDWRGLAKKLHYDRYLQFFASFPDCSPTSLL 872

Qy        861 LNLWEARHFPNGN-LSQLAAAVAGLGQPDA 889

Db        873 LDLWEASSSGSARAVPDLLQTLRVMGRPDA 902
```

RESULT 2
B44294
unc-5 protein, long form - Caenorhabditis elegans
N;Contains: unc-5 protein, short form
C;Species: Caenorhabditis elegans
C;Date: 30-Apr-1993 #sequence_revision 28-Jul-1995 #text_change 05-Nov-1999
C;Accession: B44294; T32540; A44294
R;Leung-Hagesteijn, C.; Spence, A.M.; Stern, B.D.; Zhou, Y.; Su, M.W.; Hedgecock, E.M.;
Cell 71, 289-299, 1992
A;Title: UNC-5, a transmembrane protein with immunoglobulin and thrombospondin type 1 do
A;Reference number: A44294; MUID:93046629; PMID:1384987
A;Contents: variety Bergerac
A;Accession: B44294
A;Molecule type: DNA
A;Residues: 1-947 <LEU>
A;Cross-references: GB:S47168; NID:g258527; PIDN:AAB23867.1; PID:g258529
A;Note: sequence extracted from NCBI backbone (NCBIN:116668, NCBIN:116670, NCBIN:116672,
A;Note: authors translated the codon CTA for residue 642 as Val; sequence shown follows
A;Note: mRNA lacking the first exon is equally prevalent
R;Latreille, P.
submitted to the EMBL Data Library, December 1997
A;Description: The sequence of C. elegans cosmid B0273.
A;Reference number: Z21187
A;Accession: T32540
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-947 <LAT>
A;Cross-references: EMBL:AF036698; PIDN:AAB88356.1; GSPDB:GN00022; CESP:B0273.4b
A;Experimental source: strain Bristol N2; clone B0273
C;Genetics:
A;Gene: unc-5
A;Map position: 4
A;Introns: 28/1; 69/3; 136/1; 170/3; 229/1; 351/2; 581/1; 886/3
C;Function:
A;Description: required for guidance of pioneering axons and cells migrating dorsally al
C;Superfamily: unc-5 protein; immunoglobulin homology; SH3 homology; thrombospondin type
C;Keywords: alternative splicing; duplication; glycoprotein; receptor; transmembrane pro
F;30-947/Product: unc-5 protein, short form #status predicted <ALT>
F;46-116/Domain: immunoglobulin homology <IM1>
F;153-211/Domain: immunoglobulin homology <IM2>
F;229-300/Domain: thrombospondin type 1 repeat homology #status atypical <THR1>
F;301-354/Domain: thrombospondin type 1 repeat homology <THR2>
F;365-390/Domain: transmembrane #status predicted <TMM>
F;512-559/Domain: SH3 homology <SH3>
F;53-114,65-112,160-209/Disulfide bonds: #status predicted
F;206/Binding site: carbohydrate (Asn) (covalent) #status predicted

```
  Query Match             19.9%;  Score 950.5;  DB 1;  Length 947;
  Best Local Similarity   28.5%;  Pred. No. 1.7e-61;
  Matches  265;  Conservative  165;  Mismatches  375;  Indels  125;  Gaps   31;

Qy         49 EPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQVDHVIER--STDGSSGEPTMEVRI 106

Db         37 QPKSGYVIRNKPLRLQCRANHATKIRYKCSSKWID--DSRIEKLIGTDSTSGVGYIDASV 94

Qy        107 NVSRQQVEKVFGLEEYWCQCVAWSSSG-----TTKSQKAYIRIARLRKNFEQEPLAKEVS 161

Db         95 DISRIDVDTSGHVDAFQCQCYA---SGDDDQDVVASDVATVHLAYMRKHFLKSPVAQRVQ 151

Qy        162 LEQGIVLPCRPPEGIPPAEVEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTC 221

Db        152 EGTTLQLPCQAPESDPKAELTWYKDGVVVQP--DANVIRASDGSLIMSAARLSDSGNYTC 209

Qy        222 VAKNIVARRRSASAAVIVYVNGGWSTWTEW-SVCSASCG--------------RGWQKR 265

Db        210 EATNVANSRKTDPVEVQIYVDGGWSEWSPWIGTCHVDCPLLRQHAHRIRDPHDVLPHQRR 269

Qy        266 SRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDGSWSPWSKWSACGLDCTHWRSRECSD 325

Db        270 TRTCNNPAPLNDGEYCKGEEEMTRS-CKVPCKLDGGWSSWSDWSACSSSCHRYRTRACTV 328

Qy        326 PAPRNGGEECQGTDLDTRNCTSDLCVHSASG--PEDVALYVGLIAVAVCLVLLLLVLILV 383

Db        329 PPPMNGGQPCFGDDLMTQECPAQLCTADSSRIVISDTAVYGSVASIFIVASFILAILAMF 388

Qy        384 YCR---------KKEGLDSDVADS-------------------------SILTSGFQP 407

Db        389 CCKRGNSKKSKPLKPQKMNSEKAGGIYYSEPPGVRRLLLEHQHGTLLGEKISSCSQYFEP 448

Qy        408 VSIKPSKADNPHLLTIQPDLSTTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLH 467

Db        449 PPL-------PHSTTLR---SGKSAFSGYSSTRNAGSRAALIQECSSSSSGSGGKRTMLR 498

Qy        468 HSSPT-SEAEEFVSRLSTQNYFRSLPRGTS-NMTYGTFNFLGGRLMIPNTGISLLIPPDA 525

Db        499 TSSSNCSDDDNYATLYDYMEDKSVLGLDTSQNIVAAQIDSNGARLSLSKSGARLIVPELA 558

Qy        526 IPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGV-------LLTRPVILAMDH 578

Db        559 VEGEKM--LYLAVSDTLTDQPHLKPIESALSPVIVIGQCDVSMSAHDNILRRPVVVSFRH 616

Qy        579 CGEPSP-DSWSLRLKKQSCEGSWEQDVLHLGEEAPSHLYYCQLEASA--------CYVFT 629

Db        617 CASTFPRDNWQFTL--YADEGSGWQKAVTIGEENLNTNMFVQFEQPGKKNDGFGWCHVMT 674

Qy        630 EQLGRFALVGEAL--SVAAAKRLKLLLFAPVACTSLE--YNIRVYCLHDTHDALKEVVQL 685

Db        675 YSLARLMLAGHPRRNSLSAAKRVHLAVFGPTEMSAYRRPFELRVYCVPETGAAMESVWKQ 734

Qy        686 EKQLGGQLIQEPR--VLHFKDSYHNLRLSIHDV-PSSLWKSKLLVSYQEIPFYHIWNGTQ 742

Db        735 ED--GSRLLCESNDFILNEKG---NLCICIEDVIPGFSCDGPEVVEISETQHRFV---AQ 786

Qy        743 RYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVP 802

Db        787 NGLHCSLKFRPKEINGSQFSTRVIVYQKASSTEPMVM--EVSNEPELYDATSEEREKGSV 844
```

```
Qy        803 ALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLDSHLSFFASKP--SPTAMI 860
               :      |::|| :: ::     || |     :||| ||:||| | :| |||| |   |||:::
Db        845 CV----EFRLPFGVKDELARLLDMPNESHSDWRGLAKKLHYDRYLQFFASFPDCSPTSLL 900

Qy        861 LNLWEARHFPNGN-LSQLAAAVAGLGQPDA 889
              |:||||     :    |    :   :|:|||
Db        901 LDLWEASSSGSARAVPDLLQTLRVMGRPDA 930
```

```
RESULT 3
T00026
brain-specific angiogenesis inhibitor 1 - human
N;Alternate names: BAI1 protein
C;Species: Homo sapiens (man)
C;Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 12-Feb-1999
C;Accession: T00026
R;Nishimori, H.; Shiratsuchi, T.; Urano, T.; Kimura, Y.; Kiyono, K.; Tatsumi, K.; Yoshid
submitted to the EMBL Data Library, June 1997
A;Reference number: Z14064
A;Accession: T00026
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1584 <NIS>
A;Cross-references: EMBL:AB005297; NID:d1175078; PID:d1024528
A;Experimental source: brain
C;Genetics:
A;Gene: GDB:BAI1
A;Cross-references: GDB:9838088; OMIM:602682
A;Map position: 8q24-8q24
C;Superfamily: thrombospondin type 1 repeat homology
F;408-462/Domain: thrombospondin type 1 repeat homology <THR3>

  Query Match              5.8%;  Score 280;  DB 2;  Length 1584;
  Best Local Similarity   31.8%;  Pred. No. 6.5e-12;
  Matches   74;  Conservative   38;  Mismatches   93;  Indels   28;  Gaps   10;

Qy        124 CQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAEVEW 183
              |   |     :| | |:    :|     |:    :| |  |:     |     |  | || || ||
Db        309 CNREACGPAGRTSSRSQSLRSTDARR---REELGDEL---QQFGFPA-PQTGDPAAE-EW 360

Qy        184 LRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNG 243
                 :     |   |       : ||        |     :::|:    :   :|  : :  :  :  |:|
Db        361 --SPWSVCSSTCGEGWQTR--------TRFCVSSSYSTQCSGPLREQRLCNNSAVCPVHG 410

Qy        244 GWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSSLLV---SVDG 300
                 |    |: ||:||::||||:: |:|:|     |    ||   |||      :    :    |      :|||
Db        411 AWDEWSPWSLCSSTCGRGFRDRTRTCR--PPQFGGNPCEGPEKQTKFCNIALCPGRAVDG 468

Qy        301 SWSPWSKWSACGLDCT---HWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLC 350
              :|: || ||||    |:        |:|||: |:    || ||||   ::||:|        |
Db        469 NWNEWSSWSACSASCSQGRQQRTRECNGPS--YGGAECQGHWVETRDCFLQQC 519
```

```
RESULT 4
TSHUP2
thrombospondin 2 precursor - human
C;Species: Homo sapiens (man)
C;Date: 19-May-1995 #sequence_revision 03-Aug-1995 #text_change 13-Aug-1999
C;Accession: A47379; A42173
R;LaBell, T.L.; Byers, P.H.
Genomics 17, 225-229, 1993
A;Title: Sequence and characterization of the complete human thrombospondin 2 cDNA: pote
A;Reference number: A47379; MUID:94010892; PMID:8406456
A;Accession: A47379
A;Molecule type: mRNA
A;Residues: 1-1172 <LAB>
A;Cross-references: GB:L12350; NID:g307505; PIDN:AAA03703.1; PID:g307506
R;LaBell, T.L.; Milewicz, D.J.; Disteche, C.M.; Byers, P.H.
Genomics 12, 421-429, 1992
A;Title: Thrombospondin II: partial cDNA sequence, chromosome location, and expression c
A;Reference number: A42173; MUID:92217961; PMID:1559694
A;Accession: A42173
A;Molecule type: mRNA
A;Residues: 560-1172 <LA2>
A;Cross-references: GB:M81339
A;Experimental source: fibroblast
A;Note: sequence extracted from NCBI backbone (NCBIN:95091, NCBIP:95096)
C;Genetics:
A;Gene: GDB:THBS2; TSP2
A;Cross-references: GDB:128789; OMIM:188061
A;Map position: 6q27-6q27
C;Complex: homotrimer, disulfide linked
C;Function:
A;Description: participates in cell migration and adhesion, and in platelet aggregation
C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; vc
C;Keywords: beta-hydroxyasparagine; calcium binding; cell adhesion; glycoprotein; trimer
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-1172/Product: thrombospondin 2 #status predicted <MAT>
F;319-377/Domain: von Willebrand factor type C repeat homology <VWC>
F;380-431/Domain: thrombospondin type 1 repeat homology <THR1>
F;436-492/Domain: thrombospondin type 1 repeat homology <THR2>
F;493-549/Domain: thrombospondin type 1 repeat homology <THR3>
F;553-588/Domain: EGF homology <EGF1>
F;652-691/Domain: EGF homology <EGF>
F;928-930/Region: cell attachment (R-G-D) motif
F;151,316,330,457,584,710,1069/Binding site: carbohydrate (Asn) (covalent) #status predi
F;167-226/Disulfide bonds: #status predicted
F;266,270/Disulfide bonds: interchain #status predicted
F;612/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted

  Query Match              5.7%;  Score 273.5;  DB 1;  Length 1172;
  Best Local Similarity   28.9%;  Pred. No. 1.3e-11;
  Matches   74;  Conservative   27;  Mismatches  110;  Indels   45;  Gaps    8;

Qy        209 RQARLADTANYTCVAKNIVARRRSASAA-VIVYVNGGWSTWTEWSVCSASCGRGWQKRSR 267
              :: |   |   : ||:  :|  |  ||      :||||  |: || ||  :|| :|  |  |
Db        403 QRGRSCDVTSNTCLGPSIQTRACSLSKCDTRIRQDGGWSHWSPWSSCSVTCGVGNITRIR 462

Qy        268 SCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDGSWSPWSKWSACGLDCT---HWRSRECS 324
                | :| |   || |:|       :        :|| ||||| |||| : |       |:| |:
Db        463 LCNSPVPQMGGKNCKGSGRETKACQGAPCPIDGRWSPWSPWSACTVTCAGGIRERTRVCN 522

Qy        325 DPAPRNGGEECQGTDLDTRNCTSDLCVHSASGPEDVALYVGLIAVAVCLVLLLLVLILVY 384
                | |: ||: | |    : :  |       |        | | |          | |
Db        523 SPEPQYGGKACVGDVQERQMCNKRSC------PVDGCLSNPCFPGAQC------------ 564

Qy        385 CRKKEGLDSDVADSSILTSGFQPVSI--KPSKADNPHLLTIQPDLSTTT--------TYQ 434
                        |  |  |   :  || ||     :  ::      : ||:  :|        |
Db        565 --------SSFPDGS-WSCGFCPVGFLGNGTHCEDLDECALVPDICFSTSKVPRCVNTQP 615

Qy        435 GSLC----PRQDGPSP 446
              |  |      ||  |  |
Db        616 GFHCLPCPPRYRGNQP 631
```

```
RESULT 5
A42587
thrombospondin 2 precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 20-Aug-1999
C;Accession: A42587; A39851
R;Laherty, C.D.; O'Rourke, K.; Wolf, F.W.; Katz, R.; Seldin, M.F.; Dixit, V.M.
J. Biol. Chem. 267, 3274-3281, 1992
A;Title: Characterization of mouse thrombospondin 2 sequence and expression during cell g
A;Reference number: A42587; MUID:92147683; PMID:1371115
A;Accession: A42587
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: nucleic acid
A;Residues: 1-1172 <LAH>
A;Cross-references: GB:L07803; GB:M87275; NID:g340421; PIDN:AAA53064.1; PID:g567241
A;Note: sequence extracted from NCBI backbone (NCBIP:81502)
R;Bornstein, P.; O'Rourke, K.; Wikstrom, K.; Wolf, F.W.; Katz, R.; Li, P.; Dixit, V.M.
J. Biol. Chem. 266, 12821-12824, 1991
A;Title: A second, expressed thrombospondin gene (Thbs2) exists in the mouse genome.
A;Reference number: A39851; MUID:91302287; PMID:1712771
A;Accession: A39851
```

A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-873 <BOR>
A;Cross-references: GB:M64866; NID:g201994; PIDN:AAA40432.1; PID:g201995
C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; vc
C;Keywords: calcium binding; glycoprotein
F;319-377/Domain: von Willebrand factor type C repeat homology <VWC>
F;380-431/Domain: thrombospondin type 1 repeat homology <THR1>
F;436-492/Domain: thrombospondin type 1 repeat homology <THR2>
F;493-549/Domain: thrombospondin type 1 repeat homology <THR3>
F;553-588/Domain: EGF homology <EGF1>
F;652-691/Domain: EGF homology <EGF>

```
  Query Match              5.7%;  Score 273.5;  DB 2;  Length 1172;
  Best Local Similarity   36.1%;  Pred. No. 1.3e-11;
  Matches   57;  Conservative   22;  Mismatches   70;  Indels    9;  Gaps    4;

Qy        209 RQARLADTANYTCVAKNIVARRRS-ASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSR 267
Db        403 QRGRSCDVTSNTCLGPSIQTRTCSLGKCDTRIRQNGGWSHWSPWSSCSVTCGVGNVTRIR 462

Qy        268 SCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDGSWSPWSKWSACGLDCT---HWRSRECS 324
Db        463 LCNSPVPQMGGKNCKGSGRETKPCQRDPCPIDGRWSPWSPWSACTVTCAGGIRERSRVCN 522

Qy        325 DPAPRNGGEECQG--TD---LDTRNCTSDLCVHSASGP 357
Db        523 SPEPQYGGKDCVGDVTEHQMCNKRSCPIDGCLSNPCFP 560
```

RESULT 6
JC5928
semaphorin F precursor - human
C;Species: Homo sapiens (man)
C;Date: 10-Apr-1998 #sequence_revision 08-May-1998 #text_change 17-Nov-2000
C;Accession: JC5928
R;Simmons, A.D.; Pueschel, A.W.; McPherson, J.D.; Overhauser, J.; Lovett, M.
Biochem. Biophys. Res. Commun. 242, 685-691, 1998
A;Title: Molecular cloning and mapping of human semaphorin F from the Cri-du-chat candid
A;Reference number: JC5928; MUID:98125554; PMID:9464278
A;Accession: JC5928
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-1074 <SIM>
A;Cross-references: GB:U52840; NID:g2772583; PIDN:AAC09473.1; PID:g2772584
A;Experimental source: brain
C;Comment: This protein disrupts normal brain development and leads to some of the featu
C;Genetics:
A;Gene: semaf
C;Superfamily: human semaphorin F; thrombospondin type 1 repeat homology
F;1-20/Domain: signal sequence #status predicted <SIG>
F;50-533/Domain: semaphorin #status predicted <SEM>
F;840-896/Domain: thrombospondin type 1 repeat homology <THR3>
F;971-993/Domain: transmembrane #status predicted <TMM>

```
  Query Match              5.7%;  Score 270.5;  DB 2;  Length 1074;
  Best Local Similarity   42.4%;  Pred. No. 1.9e-11;
  Matches   50;  Conservative   13;  Mismatches   52;  Indels    3;  Gaps    1;

Qy        241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDG 300
Db        783 VNGAWSAWTSWSQCSRDCSRGIRNRKRVCNNPEPKYGGMPCLGPSLEYQECNTLPCPVDG 842

Qy        301 SWSPWSKWSACGLDC---THWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCVHSAS 355
Db        843 VWSCWSPWTKCSATCGGGHYMRTRSCSNPAPAYGGDICLGLHTEEALCNTQPCPESWS 900
```

RESULT 7
T00326
hypothetical protein KIAA0550 - human
C;Species: Homo sapiens (man)
C;Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 21-Jul-2000
C;Accession: T00326
R;Nagase, T.; Ishikawa, K.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.; Ohara, O.
DNA Res. 5, 31-39, 1998
A;Title: Prediction of the coding sequences of unidentified human genes. IX. The complet
A;Reference number: Z14086; MUID:98290545; PMID:9628581
A;Accession: T00326
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-984 <NAG>
A;Cross-references: EMBL:AB011122; NID:g3043623; PIDN:BAA25476.1; PID:g3043624
A;Experimental source: brain
C;Genetics:
A;Note: KIAA0550
C;Superfamily: thrombospondin type 1 repeat homology
F;344-398/Domain: thrombospondin type 1 repeat homology <THR3>

```
  Query Match              5.6%;  Score 267.5;  DB 2;  Length 984;
  Best Local Similarity   37.7%;  Pred. No. 2.8e-11;
  Matches   55;  Conservative   22;  Mismatches   54;  Indels   15;  Gaps    5;

Qy        220 TCVA------KNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPA 273
Db        317 TCVSPYGTHCSGPLRESRVCNNTALCPVHGVWEEWSPWSLCSFTCGRGQRTRTRSCT--P 374

Qy        274 PLNGGAFCEGQNVHDRTVSSLLVSVDGSWSPWSKWSACGLDC---THWRSRECSDPAPRN 330
Db        375 PQYGGRPCEGPETHHKPCNIALCPVDGQWQEWSSWSQCSVTCSNGTQQRSRQCT--AAAH 432

Qy        331 GGEECQGTDLDTRNCTSDLCVHSASG 356
Db        433 GGSECRGPWAESRECYNPEC--TANG 456
```

RESULT 8
T00028
brain-specific angiogenesis inhibitor 3 - human
N;Alternate names: BAI3 protein
C;Species: Homo sapiens (man)
C;Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 21-Jul-2000
C;Accession: T00028
R;Shiratsuchi, T.; Nishimori, H.; Ichise, H.; Nakamura, Y.; Tokino, T.
Cytogenet. Cell Genet. 79, 103-108, 1997
A;Title: Cloning and characterization of BAI2 and BAI3, novel genes homologous to brain
A;Reference number: Z14066; MUID:98194217; PMID:9533023
A;Accession: T00028
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1522 <SHI>
A;Cross-references: EMBL:AB005299; NID:g3021700; PIDN:BAA25363.1; PID:g3021701
A;Experimental source: brain
C;Genetics:
A;Gene: GDB:BAI3
A;Cross-references: GDB:9838090; OMIM:602684
A;Map position: 6q12-6q12
C;Superfamily: thrombospondin type 1 repeat homology
F;344-398/Domain: thrombospondin type 1 repeat homology <THR3>

```
  Query Match              5.6%;  Score 267.5;  DB 2;  Length 1522;
  Best Local Similarity   37.7%;  Pred. No. 5.1e-11;
  Matches   55;  Conservative   22;  Mismatches   54;  Indels   15;  Gaps    5;

Qy        220 TCVA------KNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPA 273
Db        317 TCVSPYGTHCSGPLRESRVCNNTALCPVHGVWEEWSPWSLCSFTCGRGQRTRTRSCT--P 374

Qy        274 PLNGGAFCEGQNVHDRTVSSLLVSVDGSWSPWSKWSACGLDC---THWRSRECSDPAPRN 330
Db        375 PQYGGRPCEGPETHHKPCNIALCPVDGQWQEWSSWSQCSVTCSNGTQQRSRQCT--AAAH 432

Qy        331 GGEECQGTDLDTRNCTSDLCVHSASG 356
Db        433 GGSECRGPWAESRECYNPEC--TANG 456
```

```
RESULT 9
T18856
angiogenesis inhibitor homolog - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C;Accession: T18856; T24653
R;McMurray, A.
submitted to the EMBL Data Library, July 1995
A;Reference number: Z19031
A;Accession: T18856
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1444 <WIL>
A;Cross-references: EMBL:Z50004; PIDN:CAA90293.1; GSPDB:GN00028; CESP:C02B4.1
A;Experimental source: clone C02B4
R;McMurray, A.
submitted to the EMBL Data Library, July 1995
A;Reference number: Z19917
A;Accession: T24653
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1444 <WI2>
A;Cross-references: EMBL:Z50006; PIDN:CAA90302.1; GSPDB:GN00028; CESP:C02B4.1
A;Experimental source: clone T07C5
C;Genetics:
A;Gene: CESP:C02B4.1
A;Map position: X
A;Introns: 25/3; 70/3; 96/3; 139/3; 187/1; 234/2; 282/3; 376/2; 422/2; 478/3; 509/3; 566

  Query Match             5.5%;  Score 263.5;  DB 2;  Length 1444;
  Best Local Similarity   26.4%;  Pred. No. 9.3e-11;
  Matches   73;  Conservative   26;  Mismatches   91;  Indels   87;  Gaps   12;

Qy        123 WCQCVAWSSSGTTKSQKAYIRIARLRKNFEQ------EPLAKEVSLEQGIVLPCRPPEGI 176

Db       1134 WSEWSSWSAC-------SCFSLTSTRRRFCQVVDPTVQGFCAGAILEQ---IPCAPGSCS 1183

Qy        177 PPAE-----VEW------------LRNEDLVDPSLDPNVYITREHSLVVRQARLADTAN 218

Db       1184 PSAGGWSLWSEWSSCSKDCGDTGHQIRNRMCSEP--------------IPSNRGAYCSG 1228

Qy        219 YT-----CVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPA 273

Db       1229 YSFDQRPCVMDNVCSDEK---------VDGGWTDWTAWSECTDYCRNGHRSRTRFCANPK 1279

Qy        274 PLNGGAFCEGQN-----------VHDRTVSSLLVSVDGSWSPWSKWSACGLDC---THWR 319

Db       1280 PSQGGAQCTGSDFELNPCFDPARCHLR---------DGGWSTWSDWTPCSASCGFGVQTR 1330

Qy        320 SRECSDPAPRNGGEECQGTDLDTRNCTSDLCVHSASG 356

Db       1331 DRSCSSPEPK-GGQSCSGLAHQTSLCDLPACDHESDG 1366

RESULT 10
T00027
brain-specific angiogenesis inhibitor 2 - human
N;Alternate names: BAI2 protein
C;Species: Homo sapiens (man)
C;Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 21-Jul-2000
C;Accession: T00027
R;Shiratsuchi, T.; Nishimori, H.; Ichise, H.; Nakamura, Y.; Tokino, T.
Cytogenet. Cell Genet. 79, 103-108, 1997
A;Title: Cloning and characterization of BAI2 and BAI3, novel genes homologous to brain-
A;Reference number: Z14066; MUID:98194217; PMID:9533023
A;Accession: T00027
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1572 <SHI>
A;Cross-references: EMBL:AB005298; NID:g3021698; PIDN:BAA25362.1; PID:g3021699
A;Experimental source: brain
C;Genetics:
A;Gene: GDB:BAI2
A;Cross-references: GDB:9838089; OMIM:602683
A;Map position: 1p35-1p35

  Query Match             5.4%;  Score 258;  DB 2;  Length 1572;
  Best Local Similarity   18.8%;  Pred. No. 2.6e-10;
  Matches  172;  Conservative  110;  Mismatches  312;  Indels  322;  Gaps   37;

Qy        173 PEGIPPAEVEWLRNEDLVDPSLDPNVY------------------ITREHSLVVRQARL 213

Db        271 PEEEPKVKTQWPRSAD------EPGLYMAQTGDPAAEEWSPWSVCSLTCGQGLQVR-TRS 323

Qy        214 ADTANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPA 273

Db        324 CVSSPYGTLCSGPLRETRPCNNSATCPVHGVWEEWGSWSLCSRSCGRGSRSRMRTCV--P 381

Qy        274 PLNGGAFCEGQNVHDRTVSSLLVSVDGSWSPWSKWSACGLDC---THWRSRECSDPAPR- 329

Db        382 PQHGGKACEGPELQTKLCSMAACPVEGQWLEWGPWGPCSTSCANGTQQRSRKCSVAGPAW 441

Qy        330 ------------------------------------------------------NGGEECQ 336

Db        442 ATCTGALTDTRECSNLECPATDSKWGPWNAWSLCSKTCDTGWQRRFRMCQATGTQGYPCE 501

Qy        337 GTDLDTRNCTSDLC--VHSASGPEDVAL------------------------------- 362

Db        502 GTGEEVKPCSEKRCPAFHEMCRDEYVMLMTWKKAAAGEIIYNKCPPNASGSASRRCLLSA 561

Qy        363 ----YVGLIAVAVCL---VLLLLVLILVYCRKKEGLDSDVADSSILTSGFQPVSIKPSKA 415

Db        562 QGVAYWGLPSFARCISHEYRYLYLSLREHLAKGQRMLAGEGMSQVVRS-LQELLARRTYY 620

Qy        416 DNPHLLTIQPDLSTTTTY-QGSLCPRQDGPSPKFQLT------------NGHLLSPLGG 461

Db        621 SGDLLFSVDILRNVTDTFKRATYVPSADDVQRFFQVVSFMVDAENKEKWDDAQQVSP--G 678

Qy        462 GRHTLHHSSPTSEAEEFV-----------SRLSTQNYFRSLPRG-----TSNMTYGTFN 504

Db        679 SVHLLR------VVEDFIHLVGDALKAFQSSLIVTDNLVISIQREPVSAVSSDITFPMRG 732

Qy        505 FLG---------GRLMIPNTGISLLIP-----------------PDAIPRGK-------- 530

Db        733 RRGMKDWVRHSEDRLFLPKEVLSLSSPGKPATSGAAGSPGRGRGPGTVPPGPGHSHQRLL 792

Qy        531 ---------------IYE-IYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVIL 574

Db        793 PADPDESSYFVIGAVLYRTLGLILPPP---RPPLAVTSRVMT--VTVRPPTQPPAEPLIT 847

Qy        575 A-----MDHCGEPSPDSWSLRLKKQSCEGSWEQDVLHLGEEAPSHLYYCQLEASACYVFT 629

Db        848 VELSYIINGTTDPHCASWDYS-RADASSGDWDTENCQTLETQAAHT-RCQ---------C 896

Qy        630 EQLGRFALVGE-----ALSVAAAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQ 684

Db        897 QHLSTFAVLAQPPKDLTLELAGSPSVPLVIGCAVSCMALLTLLAIYA------AFWRFIK 950

Qy        685 LEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWKSKLLVSYQEIPFYHIWNGTQRY 744

Db        951 SERSI--------------ILLNFCLSI--LASNI---LILVGQSRVLSKGVCTMTAAF 990

Qy        745 LHCTFTLERVSPSTSDLACKLWV--------------------------WQVEGDGQS 776

Db        991 LHFFF-----------LSSFCWVLTEAWQSYLAVIGRMRTRLVRKRFLCLGWGLPALVVA 1039

Qy        777 FSINFNITKDTRFAELLALESEAG-VPALVGPSA------FKIPFLIRQKI-----ISSL 824

Db       1040 VSVGFTRTKGYGTSSYCWLSLEGGLLYAFVGPAAVIVLVNMLIGIIVFNKLMARDGISDK 1099

Qy        825 DPPCRRGAD---WRTL 837

Db       1100 SKKQRAGSERCPWASL 1115
```

RESULT 11
A40558
thrombospondin 1 precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 05-Jun-1992 #sequence_revision 05-Jun-1992 #text_change 20-Aug-1999
C;Accession: A40558; A37905; B42587; S68787
R;Lawler, J.; Duquette, M.; Ferro, P.; Copeland, N.G.; Gilbert, D.J.; Jenkins, N.A.
Genomics 11, 587-600, 1991
A;Title: Characterization of the murine thrombospondin gene.
A;Reference number: A40558; MUID:92128941; PMID:1774063
A;Accession: A40558
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1170 <LAW>
A;Cross-references: GB:M62449; GB:M62450; GB:M62451; GB:M62452; GB:M62453; GB:M62454; GB
M62465; GB:M62466; GB:M62467; GB:M62468; GB:M62469; GB:M62470; NID:g511867; PIDN:AAA5061
R;Bornstein, P.; Alfi, D.; Devarayalu, S.; Framson, P.; Li, P.
J. Biol. Chem. 265, 16691-16698, 1990
A;Title: Characterization of the mouse thrombospondin gene and evaluation of the role of
A;Reference number: A37905; MUID:90375546; PMID:2398070
A;Accession: A37905
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-490 <BOR>
A;Cross-references: GB:J05605; GB:J05606; NID:g201991; PIDN:AAA40431.1; PID:g554390
R;Laherty, C.D.; O'Rourke, K.; Wolf, F.W.; Katz, R.; Seldin, M.F.; Dixit, V.M.
J. Biol. Chem. 267, 3274-3281, 1992
A;Title: Characterization of mouse thrombospondin 2 sequence and expression during cell
A;Reference number: A42587; MUID:92147683; PMID:1371115
A;Accession: B42587
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-1152,'P',1154-1170 <LAH>
A;Cross-references: GB:M87276
A;Note: sequence extracted from NCBI backbone (NCBIP:81501)
R;Chen, H.; Aeschlimann, D.; Nowlen, J.; Mosher, D.F.
FEBS Lett. 387, 36-41, 1996
A;Title: Expression and initial characterization of recombinant mouse thrombospondin 1 a
A;Reference number: S68787; MUID:96234006; PMID:8654563
A;Accession: S68787
A;Molecule type: protein
A;Residues: 19-26,'X',28-37 <CHE>
C;Complex: homotrimer, disulfide linked
C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; vo
C;Keywords: calcium binding; glycoprotein; homotrimer
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-1170/Product: thrombospondin 1 #status predicted <MAT>
F;317-375/Domain: von Willebrand factor type C repeat homology <VWC>
F;378-429/Domain: thrombospondin type 1 repeat homology <THR1>
F;434-490/Domain: thrombospondin type 1 repeat homology <THR2>
F;491-547/Domain: thrombospondin type 1 repeat homology <THR3>
F;551-586/Domain: EGF homology <EGF>
F;248,360,708,1067/Binding site: carbohydrate (Asn) (covalent) #status predicted

```
  Query Match             5.2%;  Score 248;  DB 2;  Length 1170;
  Best Local Similarity   29.9%;  Pred. No. 9.5e-10;
  Matches   53;  Conservative   24;  Mismatches   76;  Indels   24;  Gaps    4;

Qy        207 VVRQARLADTANYTCVAKNIVAR--------RRSASAAVIVYVNGGWSTWTEWSVCSASC 258
              : ::  |  |:  |  |    ::  |         :|           :||||  |:  ||  ||  :|
Db        399 IQQRGRSCDSLNNRCEGSSVQTRTCHIQECDKRFKQ-------DGGWSHWSPWSSCSVTC 451

Qy        259 GRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDGSWSPWSKWSACGLDC--- 315
              | |     | |  |  :|:|     |    |||:        :        ::|  |  |||  |    |  :  |
Db        452 GDGVITRIRLCNSPSPQMNGKPCEGEARETKACKKDACPINGGWGPWSPWDICSVTCGGG 511

Qy        316 THWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCVHSASGPEDVALYVGLIAVAVC 372
                    ||| |::|  |:  ||::|  |      :   :   |           |      |      |       |  |  |
Db        512 VQRRSRLCNNPTPQFGGKDCVGDVTENQVCNKQDC------PIDGCLSNPCFAGAKC 562
```

RESULT 12
TSHUP1
thrombospondin 1 precursor - human
C;Species: Homo sapiens (man)
C;Date: 23-Aug-1987 #sequence_revision 03-Aug-1995 #text_change 17-Nov-2000
C;Accession: A26155; A34274; A30140; A25812; A05172; A42927
R;Lawler, J.; Hynes, R.O.
J. Cell Biol. 103, 1635-1648, 1986
A;Title: The structure of human thrombospondin, an adhesive glycoprotein with multiple c
A;Reference number: A26155; MUID:87057617; PMID:2430973
A;Accession: A26155
A;Molecule type: mRNA
A;Residues: 1-1170 <LAW>
A;Cross-references: GB:X04665; NID:g37137; PIDN:CAA28370.1; PID:g37138
A;Note: parts of this sequence, including the amino end of the mature protein, were dete
R;Laherty, C.D.; Gierman, T.M.; Dixit, V.M.
J. Biol. Chem. 264, 11222-11227, 1989
A;Title: Characterization of the promoter region of the human thrombospondin gene. DNA s
A;Reference number: A34274; MUID:89291870; PMID:2544587
A;Accession: A34274
A;Molecule type: DNA
A;Residues: 1-166 <LAH>
A;Cross-references: GB:J04835
R;Hennessy, S.W.; Frazier, B.A.; Kim, D.D.; Deckwerth, T.L.; Baumgartel, D.M.; Rotwein,
J. Cell Biol. 108, 729-736, 1989
A;Title: Complete thrombospondin mRNA sequence includes potential regulatory sites in th
A;Reference number: A30140; MUID:89139590; PMID:2918029
A;Accession: A30140
A;Molecule type: mRNA
A;Residues: 1-83,'A',85-522,'A',524-1170 <HEN>
A;Cross-references: EMBL:X14787; NID:g37464; PIDN:CAA32889.1; PID:g37465
A;Note: parts of this sequence, including the amino end of the mature protein, were dete
R;Kobayashi, S.; Eden-McCutchan, F.; Framson, P.; Bornstein, P.
Biochemistry 25, 8418-8425, 1986
A;Title: Partial amino acid sequence of human thrombospondin as determined by analysis o
A;Reference number: A25812; MUID:87157592; PMID:3030396
A;Accession: A25812
A;Molecule type: mRNA
A;Residues: 1-83,'A',85-397 <KOB>
A;Cross-references: GB:M25631; NID:g538353; PIDN:AAA36741.1; PID:g538354
R;Dixit, V.M.; Hennessy, S.W.; Grant, G.A.; Rotwein, P.; Frazier, W.A.
Proc. Natl. Acad. Sci. U.S.A. 83, 5449-5453, 1986
A;Reference number: A05172; MUID:86287276; PMID:3461443
A;Accession: A05172
A;Molecule type: mRNA
A;Residues: 1-83,'A',85-374,'RC' <DIX>
A;Cross-references: GB:M14326; NID:g340005; PIDN:AAA61237.1; PID:g553801
A;Note: parts of this sequence, including the amino end of the mature protein, were dete
R;Sun, X.; Skorstengaard, K.; Mosher, D.F.
J. Cell Biol. 118, 693-701, 1992
A;Title: Disulfides modulate RGD-inhibitable cell adhesive activity of thrombospondin.
A;Reference number: A42927; MUID:92348511; PMID:1379247
A;Accession: A42927
A;Molecule type: protein
A;Residues: 987-1003 <SUN>
A;Note: Cys-992 is shown to have a free sulfhydryl
C;Genetics:
A;Gene: GDB:THBS1; TSP1; TSP
A;Cross-references: GDB:120438; OMIM:188060
A;Map position: 15q15-15q15
A;Introns: 23/1
A;Note: the list of introns may be incomplete
C;Complex: homotrimer, disulfide linked
C;Function:
A;Description: participates in cell migration and adhesion, and in platelet aggregation
C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; vo
C;Keywords: beta-hydroxyasparagine; calcium binding; cell adhesion; glycoprotein; trimer
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-1170/Product: thrombospondin 1 #status predicted <MAT>
F;317-375/Domain: von Willebrand factor type C repeat homology <VWC>
F;378-429/Domain: thrombospondin type 1 repeat homology <THR1>
F;434-490/Domain: thrombospondin type 1 repeat homology <THR2>
F;491-547/Domain: thrombospondin type 1 repeat homology <THR3>
F;551-586/Domain: EGF homology <EGF1>
F;650-689/Domain: EGF homology <EGF2>

```
F;926-928/Region: cell attachment (R-G-D) motif
F;171-232/Disulfide bonds: #status predicted
F;248,360,708,1067/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;270,274/Disulfide bonds: interchain #status predicted
F;610/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted
F;1051/Binding site: carbohydrate (Asn) (covalent) #status absent

  Query Match             5.1%;  Score 246;  DB 1;  Length 1170;
  Best Local Similarity   30.3%;  Pred. No. 1.3e-09;
  Matches   47;  Conservative   24;  Mismatches   66;  Indels   18;  Gaps    3;

Qy        207 VVRQARLADTANYTCVAKNIVAR--------RRSASAAVIVYVNGGWSTWTEWSVCSASC 258
              : ::  |  |:  |   |    ::  |            :|         :||||  |: || || :|
Db        399 IQQRGRSCDSLNNRCEGSSVQTRTCHIQECDKRFKQ-------DGGWSHWSPWSSCSVTC 451

Qy        259 GRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDGSWSPWSKWSACGLDC--- 315
              | |    | | |  :|:|    |  |||:     :            ::| |  ||| |   |  : |
Db        452 GDGVITRIRLCNSPSPQMNGKPCEGEARETKACKKDACPINGGWGPWSPWDICSVTCGGG 511

Qy        316 THWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLC 350
                  ||| |::|  |:  ||::|  |     :  :  |       |
Db        512 VQKRSRLCNNPTPQFGGKDCVGDVTENQICNKQDC 546
```

```
RESULT 13
A39804
thrombospondin precursor - chicken
C;Species: Gallus gallus (chicken)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A39804
R;Lawler, J.; Duquette, M.; Ferro, P.
J. Biol. Chem. 266, 8039-8043, 1991
A;Title: Cloning and sequencing of chicken thrombospondin.
A;Reference number: A39804; MUID:91217026; PMID:2022631
A;Accession: A39804
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1178 &lt;LAW&gt;
A;Cross-references: GB:M60853; NID:g212763; PIDN:AAA51437.1; PID:g212764
C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; vo
F;325-383/Domain: von Willebrand factor type C repeat homology &lt;VWC&gt;
F;386-437/Domain: thrombospondin type 1 repeat homology &lt;THR1&gt;
F;442-498/Domain: thrombospondin type 1 repeat homology &lt;THR2&gt;
F;499-555/Domain: thrombospondin type 1 repeat homology &lt;THR3&gt;
F;658-697/Domain: EGF homology &lt;EGF&gt;

  Query Match             5.1%;  Score 243.5;  DB 1;  Length 1178;
  Best Local Similarity   34.4%;  Pred. No. 2.1e-09;
  Matches   55;  Conservative   15;  Mismatches   75;  Indels   15;  Gaps    4;

Qy        210 QARLADTANYTCVAKNIVARRRS-ASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRS 268
               :  |   |      |   :|  |   |           :   :||||  |:  || ||  :||  |     |  |
Db        410 RGRSCDVTRSACTGPHIQTRMCSFKKCDHRIRQDGGWSHWSPWSSCSVTCGVGNITRIRL 469

Qy        269 CTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDGSWSPWSKWSACGLDC---THWRSRECSD 325
              | :|  |   ||  | |                |:|  |  |||  ||||  :  |         ||| |:
Db        470 CNSPIPQMGGKNCVGNGRETEKCEKAPCPVNGQWGPWSPWSACTVTCGGGIRERSRLCNS 529

Qy        326 PAPRNGGEECQGTDLDT---------RNCTSDLCVHSASGP 357
              | |:  ||:  |  |    ||           |:|   |  |:  :     |
Db        530 PEPQYGGKPCVG---DTKQHDMCNKRDCPIDGCLSNPCFP 566
```

```
RESULT 14
T14160
transmembrane receptor protein Robo1 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C;Accession: T14160
R;Kidd, T.; Brose, K.; Mitchell, K.J.; Fetter, R.D.; Tessier-Lavigne, M.; Goodman, C.S.;
Cell 92, 205-215, 1998
A;Title: Roundabout controls axon crossing of the CNS midline and defines a novel subfam
A;Reference number: Z17897; MUID:98117249; PMID:9458045
```

```
A;Accession: T14160
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1651 &lt;KID&gt;
A;Cross-references: EMBL:AF041082; NID:g2811215; PID:g2811216; PIDN:AAC39960.1
C;Function:
A;Description: appears to function as the gatekeeper controlling midline crossing
C;Keywords: transmembrane protein

  Query Match             4.6%;  Score 220.5;  DB 2;  Length 1651;
  Best Local Similarity   20.1%;  Pred. No. 1.6e-07;
  Matches  205;  Conservative  132;  Mismatches  377;  Indels  305;  Gaps   48;

Qy         41 DLLPHFLVEPEDVYIVKNKPVLLVCKAV--PATQIFFKCNGEWVRQVDHVIERS------ 92
              |   |   :   |  |: :  | :|  |  |||   |     |  :     || |  :  |      ||
Db         65 DFPPRIVEHPSDLIVSKGEPATLNCKAEGRPTPTIEWYKGGERV-ETDKDDPRSHRMLLP 123

Qy         93 ---------TDGSSGEPTMEVRINVSRQQVEKVFGLEEYWCQCVAWSSSGTTKSQKAYIR 143
                       |       |     | |                      ||| :    |       |   | :
Db        124 SGSLFFLRIVHGRKSRPDEGVYI------------------CVARNYLGEAVSHNASLE 164

Qy        144 IARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAEVEWLRNEDLVDPSLDPNVYIT-R 202
              :| || :|  |  |     |:: :    |:  |:||  |      :  |   :|    |  :  || |
Db        165 VAILRDDFRQNPSDVMVAVGEPAVMECQPPRGHPEPTISWKKDGSPLD---DKDERITIR 221

Qy        203 EHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNGGW-------STWTEWSV-- 253
                |::    | :|   | ||  |:|   |   |   |  |     :            :    :  |
Db        222 GGKLMITYTRKSDAGKYVCVGTNMVGERESKVADVTVLERPSFVKRPSNLAVTVDDSAEF 281

Qy        254 -CSASCGR----GWQK-----------------RSRSCTNPAPLNGGAFCEGQNVHDRT 290
               | |          ||:|                    : |   |   |    |     |  :|:  :
Db        282 KCEARGDPVPTFGWRKDDGELPKSRYEIRDDHTLKIRKVT--AGDMGSYTCVAENMVGKA 339

Qy        291 VSSLLVSVDGS----WSPWSKWSACGLDCTHWRSRECSDPAP-----RNGGEEC------ 335
               :|  ::|           |   :  |  |    | ::       :|  |        | |  :
Db        340 EASATLTVQEPPHFVVKPRDQVVALGRTVT-FQCEATGNPQPAIFWRREGSQNLLFSYQP 398

Qy        336 -QGTDLDTRNCTSDLCVHSASGPEDVALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSD 394
               |  :    :  :  | || |  :               ||  |           :|  |                  :
Db        399 PQSSSRFSVSQTGDLTVTNVQ-RSDVGYY-------ICQTL------------------N 432

Qy        395 VADSSILTSGFQPVSIKPSKADNPHLLTIQPDLSTTTTYQGSL---CPRQDGPSPKF--- 448
              ||  ||:|   :   :  :     || |    :   |    :: |      |:|     |       | |
Db        433 VA-GSIITKAY--LEVTDVIADRPPPVIRQGPVNQTVAVDGTLTLSCVATGSPVPTILWR 489

Qy        449 -------------QLTNGHL---LSPLGG-GRHTLHHSSPTSEAEEFVSRLSTQNYFRS 490
                            || :| |     :  ||   ||:|    |:|: ||
Db        490 KDGVLVSTQDSRIKQLESGVLQIRYAKLGDTGRYTCTASTPSGEA-------------- 534

Qy        491 LPRGTSNMTYGTFNFLGGRLMIPNTGISLLIP----PDAIPRGKIYEIYLTLHKPEDVRL 546
                          |:   :          :  :    |: :   |      |: ||                 |||   :
Db        535 --------TWSAY------IEVQEFGVPVQPPRPTDPNLIPSAP--------SKPEVTDV 572

Qy        547 PLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGSWEQDVLH 606
                            |    | ::         |       |  :|  |     |     ||             |:    :   :   :
Db        573 SKNTVTLLWQPNLN---SGATPTSYIIEAFSHA---SGSSW------QTVAENVKTETFA 620

Qy        607 LGEEAPSHLYYCQLEASACYVFTE---------------------------QLGRFALVG 639
               :      |: :|     :  |:    |   ::                                :||
Db        621 IKGLKPNAIYLFLVRAANAYGISDPSQISDPVKTQDVPPTTQGVDHKQVQRELG------ 674

Qy        640 EALSVAAAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRV 699
                              :  |   |       |   :|        :           |  |  |       :  :|            |  :::  |
Db        675 ---------NVVLHLHNPTILSSSSVEV-----HWTVDQQSQYIQ-----GYKILYRPS- 714

Qy        700 LHFKDSYHNLRLSIHDVPSSLWKSKLL------VSYQ--EIPFYHIWNGTQRYLHCTFTL 751
                  |:       :  :|  :       |  ::         |:|:         ||:: :  |       :       ||
Db        715 ---GASHGESEWLVFEVRTPTKNSVVIPDLRKGVNYEIKARPFFNEFQGADSEIKFAKTL 771

Qy        752 ERVSPSTSDLACKLWVWQVEGDGQSFSINFN-ITKDTRFAEL-------LALESEAGVPA 803
              |     ||           :     :   |  :  :|:|  :    :   :          :||:        :             |  |:        :
Db        772 EE-RPSAPPRS--VTVSKNDGNGTAILVTWQPPPEDTQNGMVQEYKVWCLGNETRYHINK 828
```

```
Qy        804 LVGPSAFK--IPFLIRQKIISSLDPPCRRGA--DWRTLAQKLHLDSH------------ 846

Db        829 TVDGSTFSVVIPFLV-PGIRYSVEVAASTGAGPGVKSEPQFIQLDSHGNPVSPEDQVSLA 887

Qy        847 --LSFFASKPSPTA----------MILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFT 893

Db        888 QQISDVVKQPAFIAGIGAACWIILMVFSIWLYRHRKKRN--GLSSTYAGIRKVPSFTFT 944
```

```
RESULT 15
T30805
dutt1 protein - mouse
N;Alternate names: transmembrane receptor protein Robo1 homolog
C;Species: Mus musculus (house mouse)
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C;Accession: T30805
R;Wu, M.C.; Lowe, N.; Fordham, R.; Rabbitts, P.
submitted to the EMBL Data Library, July 1998
A;Description: The mouse homologue of human DUTT1/H-robo1 gene: protein sequence and chr
A;Reference number: Z20879
A;Accession: T30805
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1612 <WUM>
A;Cross-references: EMBL:Y17793; NID:e1329712; PID:e1329713; PIDN:CAA76850.1
A;Experimental source: brain
C;Genetics:
A;Gene: dutt1
A;Map position: 16

  Query Match              4.5%;  Score 217.5;  DB 2;  Length 1612;
  Best Local Similarity   19.9%;  Pred. No. 2.6e-07;
  Matches  211;  Conservative  133;  Mismatches  359;  Indels  355;  Gaps   52;

Qy         41 DLLPHFLVEPEDVYIVKNKPVLLVCKAV--PATQIFFKCNGEWVRQVDHVIERS------ 92

Db         26 DFPPRIVEHPSDLIVSKGEPATLNCKAEGRPTPTIEWYKGGERV-ETDKDDPRSHRMLLP 84

Qy         93 ---------TDGSSGEPTMEVRINVSRQQVEKVFGLEEYWCQCVAWSSSGTTKSQKAYIR 143

Db         85 SGSLFFLRIVHGRKSRPDEGVYI-------------------CVARNYLGEAVSHNASLE 125

Qy        144 IARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAEVEWLRNEDLVDPSLDPNVYIT-R 202

Db        126 VAILRDDFRQNPSDVMVAVGEPAVMECQPPRGHPEPTISWKKDGSPLD---DKDERITIR 182

Qy        203 EHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNGGW-------STWTEWSV-- 253

Db        183 GGKLMITYTRKSDAGKYVCVGTNMVGERESEVAELTVLERPSFVKRPSNLAVTVDDSAEF 242

Qy        254 -CSASCGRG-------WQK-----------------RSRSCTNPAPLNGGAFCEGQNVH 287

Db        243 KCEA---RGDPVPTVRWRKDDGELPKSRYEIRDDHTLKIRKVT--AGDMGSYTCVAENMV 297

Qy        288 DRTVSSLLVSVDGS----WSPWSKWSACGLDCTHWRSRECSDPAP-----RNGGEEC--- 335

Db        298 GKAEASATLTVQEPPHFVVKPRDQVVALGRTVT-FQCEATGNPQPAIFWRREGSQNLLFS 356

Qy        336 ----QGTDLDTRNCTSDLCVHSASGPEDVALYVGLIAVAVCLVLLLLVLILVYCRKKEGL 391

Db        357 YQPPQSSSRFSVSQTGDLTITNVQ-RSDVGYY-------ICQTL---------------- 392

Qy        392 DSDVADSSILTSGFQPVSIKPSKADNPHLLTIQPDLSTTTTYQGSL---CPRQDGPSPKF 448

Db        393 --NVA-GSIITKAY--LEVTDVIADRPPPVIRQGPVNQTVAVDGTLILSCVATGSPAPTI 447

Qy        449 -----------------QLTNGHL---LSPLGG-GRHTLHHSSPTSEAEEFVSRLSTQNY 487

Db        448 LWRKDGVLVSTQDSRIKQLESGVLQIRYAKLGDTGRYTCTASTPSGEA------------ 495

Qy        488 FRSLPRGTSNMTYGTFNFLGGRLMIPNTGISLLIP----PDAIPRGKIYEIYLTLHKPE- 542

Db        496 -----------TWSAY------IEVQEFGVPVQPPRPTDPNLIPSAP--------SKPEV 530

Qy        543 -DV---RLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEG 598

Db        531 TDVSKNTVTLSWQPNLNS--------GATPTSYIIEAFSHA---SGSSW------QTAAE 573

Qy        599 SWEQDVLHLGEEAPSHLYYCQLEASACYVFTE--------------------------Q 631

Db        574 NVKTETFAIKGLKPNAIYLFLVRAANAYGISDPSQISDPVKTQDVPPTSQGVDHKQVQRE 633

Qy        632 LGRFALVGEALSVAAAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGG 691

Db        634 LG--------------NVVLHLHNPTILSSSSVEV-----HWTVDQQSQYIQ-----GY 668

Qy        692 QLIQEPRVLHFKDSYHNLRLSIHDVPSSLWKSKLL------VSYQ--EIPFYHIWNGTQR 743

Db        669 KILYRPS----GASHGESEWLVFEVRTPTKNSVVIPDLRKGVNYEIKARPFFNEFQGADS 724

Qy        744 YLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFN-ITKDTRFAEL-------LAL 795

Db        725 EIKFAKTLEE-APSAPPRS--VTVSKNDGNGTAILVTWQPPPEDTQNGMVQEYKVWCLGN 781

Qy        796 ESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCR----------RGADWRTLAQKLHLDS 845

Db        782 ETKYHINKTVDGSTFSV-------VIPSLVPGIRYSVEVAASTGAGPGVKSEPQFIQLDS 834

Qy        846 H---------------LSFFASKPSPTA----------MILNLWEARHFP--NGNLSQLA 878

Db        835 HGNPVSPEDQVSLAQQISDVVRQPAFIAGIGAACWIILMVFSIWLYRHRKKRNGLTSTYA 894

Qy        879 A--------------------AVAGLGQPDAGLFTVSE 896

Db        895 GIRKVPSFTFTPTVTYQRGGEAVSSGGRP--GLLNISE 930

Search completed: October  6, 2004, 18:20:08
Job time : 33 secs
```

GenCore version 5.1.6


```
OM protein - protein search, using sw model

Run on:          October  6, 2004, 15:50:42 ; Search time 19 Seconds
                                             (without alignments)
                                             2463.740 Million cell updates/sec

Title:           US-09-970-944-2
Perfect score:   4787
Sequence:        1 MAVRPGLWPALLGIVLAAWL..........AVAGLGQPDAGLFTVSEAEC 899

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters:      141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SwissProt_42:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.
```

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 280 | 5.8 | 1584 | 1 | BAI1_HUMAN | O14514 homo sapien |
| 2 | 273.5 | 5.7 | 1172 | 1 | TSP2_HUMAN | P35442 homo sapien |
| 3 | 273.5 | 5.7 | 1172 | 1 | TSP2_MOUSE | Q03350 mus musculu |
| 4 | 272.5 | 5.7 | 1077 | 1 | SM5A_MOUSE | Q62217 mus musculu |
| 5 | 270.5 | 5.7 | 1074 | 1 | SM5A_HUMAN | Q13591 homo sapien |
| 6 | 269 | 5.6 | 1170 | 1 | TSP2_BOVIN | Q95116 bos taurus |
| 7 | 267.5 | 5.6 | 1522 | 1 | BAI3_HUMAN | O60242 homo sapien |
| 8 | 266.5 | 5.6 | 1173 | 1 | TSP1_XENLA | P35448 xenopus lae |
| 9 | 258.5 | 5.4 | 1093 | 1 | SM5B_MOUSE | Q60519 mus musculu |
| 10 | 258 | 5.4 | 1572 | 1 | BAI2_HUMAN | O60241 homo sapien |
| 11 | 252 | 5.3 | 1093 | 1 | SM5B_HUMAN | Q9p283 homo sapien |
| 12 | 248 | 5.2 | 1170 | 1 | TSP1_MOUSE | P35441 mus musculu |
| 13 | 246 | 5.1 | 1170 | 1 | TSP1_HUMAN | P07996 homo sapien |
| 14 | 243.5 | 5.1 | 1178 | 1 | TSP2_CHICK | P35440 gallus gall |
| 15 | 243 | 5.1 | 1170 | 1 | TSP1_BOVIN | Q28178 bos taurus |
| 16 | 208.5 | 4.4 | 469 | 1 | PROP_HUMAN | P27918 homo sapien |
| 17 | 205.5 | 4.3 | 470 | 1 | PROP_CAVPO | Q64181 cavia porce |
| 18 | 202.5 | 4.2 | 437 | 1 | PROP_MOUSE | P11680 mus musculu |
| 19 | 186 | 3.9 | 1266 | 1 | NGCA_CHICK | Q03696 gallus gall |
| 20 | 181.5 | 3.8 | 867 | 1 | SSPO_BOVIN | P98167 bos taurus |
| 21 | 178.5 | 3.7 | 1223 | 1 | AT14_HUMAN | Q8wxs8 homo sapien |
| 22 | 175.5 | 3.7 | 1736 | 1 | ZO1_HUMAN | Q07157 homo sapien |
| 23 | 172 | 3.6 | 630 | 1 | ATS4_RAT | Q9esp7 rattus norv |
| 24 | 170.5 | 3.6 | 1745 | 1 | ZO1_MOUSE | P39447 mus musculu |
| 25 | 169 | 3.5 | 934 | 1 | CO6_HUMAN | P13671 homo sapien |
| 26 | 167.5 | 3.5 | 905 | 1 | ATS8_MOUSE | P57110 mus musculu |
| 27 | 166 | 3.5 | 837 | 1 | ATS4_HUMAN | O75173 homo sapien |
| 28 | 162.5 | 3.4 | 587 | 1 | CO8B_ONCMY | Q90x85 oncorhynchu |
| 29 | 160.5 | 3.4 | 562 | 1 | AT15_MOUSE | P59384 mus musculu |
| 30 | 160.5 | 3.4 | 967 | 1 | ATS1_HUMAN | Q9uhi8 homo sapien |
| 31 | 160.5 | 3.4 | 1077 | 1 | AT10_HUMAN | Q9h324 homo sapien |
| 32 | 159.5 | 3.3 | 1224 | 1 | AT16_HUMAN | Q8te57 homo sapien |
| 33 | 156.5 | 3.3 | 860 | 1 | ATS6_HUMAN | Q9ukp5 homo sapien |
| 34 | 156 | 3.3 | 950 | 1 | AT15_HUMAN | Q8te58 homo sapien |
| 35 | 155 | 3.2 | 997 | 1 | ATS7_HUMAN | Q9ukp4 homo sapien |
| 36 | 154.5 | 3.2 | 1593 | 1 | AT12_HUMAN | P58397 homo sapien |
| 37 | 154 | 3.2 | 1213 | 1 | ATS2_MOUSE | Q8c9w3 m adamts-2 |
| 38 | 153.5 | 3.2 | 1205 | 1 | ATS3_HUMAN | O15072 homo sapien |
| 39 | 152 | 3.2 | 967 | 1 | ATS1_RAT | Q9wuq1 rattus norv |
| 40 | 152 | 3.2 | 968 | 1 | ATS1_MOUSE | P97857 mus musculu |
| 41 | 151.5 | 3.2 | 890 | 1 | ATS8_HUMAN | Q9up79 homo sapien |
| 42 | 151.5 | 3.2 | 1095 | 1 | AT17_HUMAN | Q8te56 homo sapien |
| 43 | 151 | 3.2 | 930 | 1 | ATS5_HUMAN | Q9una0 homo sapien |
| 44 | 150 | 3.1 | 1277 | 1 | CAML_FUGRU | Q98902 fugu rubrip |
| 45 | 149 | 3.1 | 140 | 1 | CO8B_RAT | P55314 rattus norv |

ALIGNMENTS

```
RESULT 1
BAI1_HUMAN
ID   BAI1_HUMAN     STANDARD;      PRT;  1584 AA.
AC   O14514;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Brain-specific angiogenesis inhibitor 1 precursor.
GN   BAI1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fetal brain;
RX   MEDLINE=98054121; PubMed=9393972;
RA   Nishimori H., Shiratsuchi T., Urano T., Kimura Y., Kiyono K.,
RA   Tatsumi K., Yoshida S., Ono M., Kuwano M., Nakamura Y., Tokino T.;
RT   "A novel brain-specific p53-target gene, BAI1, containing
RT   thrombospondin type 1 repeats inhibits experimental angiogenesis.";
RL   Oncogene 15:2145-2150(1997).
RN   [2]
RP   INTERACTION WITH BAP1.
RX   MEDLINE=98321173; PubMed=9647739;
RA   Shiratsuchi T., Futamura M., Oda K., Nishimori H., Nakamura Y.,
RA   Tokino T.;
RT   "Cloning and characterization of BAI-associated protein 1: a PDZ
RT   domain-containing protein that interacts with BAI1.";
RL   Biochem. Biophys. Res. Commun. 247:597-604(1998).
CC   -!- FUNCTION: LIKELY TO BE A POTENT INHIBITOR OF ANGIOGENESIS IN
CC       BRAIN AND MAY PLAY A SIGNIFICANT ROLE AS A MEDIATOR OF THE P53
CC       SIGNAL IN SUPPRESSION OF GLIOBLASTOMA. MAY FUNCTION IN CELL
CC       ADHESION AND SIGNAL TRANSDUCTION IN THE BRAIN.
CC   -!- SUBUNIT: INTERACTS WITH BAP1.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. LIKELY TO BE
CC       CONCENTRATED AT CELL-CELL ADHESION SITES.
CC   -!- TISSUE SPECIFICITY: SPECIFICALLY EXPRESSED IN BRAIN. REDUCED OR NO
CC       EXPRESSION IS OBSERVED IN SOME GLIOBLASTOMA CELL LINES AND CANCER
CC       TISSUES.
CC   -!- INDUCTION: By p53.
CC   -!- DOMAIN: THE TSP1 REPEATS INHIBIT IN VIVO ANGIOGENESIS IN RAT
CC       CORNEA INDUCED BY BFGF.
CC   -!- SIMILARITY: Belongs to family 2 of G-protein coupled receptors.
CC   -!- SIMILARITY: Contains 5 TSP type-1 domains.
CC   -!- SIMILARITY: Contains 1 GPS domain.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AB005297; BAA23647.1; -.
```

```
DR   PIR; T00026; T00026.
DR   Genew; HGNC:943; BAI1.
DR   MIM; 602682; -.
DR   GO; GO:0005887; C:integral to plasma membrane; TAS.
DR   GO; GO:0005911; C:intercellular junction; TAS.
DR   GO; GO:0005515; F:protein binding; TAS.
DR   GO; GO:0007409; P:axonogenesis; TAS.
DR   GO; GO:0007155; P:cell adhesion; TAS.
DR   GO; GO:0008285; P:negative regulation of cell proliferation; TAS.
DR   GO; GO:0007422; P:peripheral nervous system development; TAS.
DR   GO; GO:0007165; P:signal transduction; TAS.
DR   InterPro; IPR000832; GPCR_secretin.
DR   InterPro; IPR001879; hormn_receptor.
DR   InterPro; IPR000203; PKD_cys_rich.
DR   InterPro; IPR000884; TSP1.
DR   Pfam; PF00002; 7tm_2; 1.
DR   Pfam; PF01825; GPS; 1.
DR   Pfam; PF02793; HRM; 1.
DR   Pfam; PF00090; tsp_1; 5.
DR   SMART; SM00303; GPS; 1.
DR   SMART; SM00008; HormR; 1.
DR   SMART; SM00209; TSP1; 5.
DR   PROSITE; PS50221; GPS; 1.
DR   PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; FALSE_NEG.
DR   PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; FALSE_NEG.
DR   PROSITE; PS50227; G_PROTEIN_RECEP_F2_3; 1.
DR   PROSITE; PS50261; G_PROTEIN_RECEP_F2_4; 1.
DR   PROSITE; PS50092; TSP1; 5.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW   Repeat; Cell adhesion.
FT   SIGNAL        1     30       POTENTIAL.
FT   CHAIN        31   1584       BRAIN-SPECIFIC ANGIOGENESIS INHIBITOR 1.
FT   DOMAIN       31    948       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    949    969       1 (POTENTIAL).
FT   DOMAIN      970    980       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    981   1001       2 (POTENTIAL).
FT   DOMAIN     1002   1008       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   1009   1029       3 (POTENTIAL).
FT   DOMAIN     1030   1052       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   1053   1073       4 (POTENTIAL).
FT   DOMAIN     1074   1093       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   1094   1114       5 (POTENTIAL).
FT   DOMAIN     1115   1136       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   1137   1157       6 (POTENTIAL).
FT   DOMAIN     1158   1166       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   1167   1187       7 (POTENTIAL).
FT   DOMAIN     1188   1584       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      261    315       TSP TYPE-1 1.
FT   DOMAIN      354    407       TSP TYPE-1 2.
FT   DOMAIN      409    462       TSP TYPE-1 3.
FT   DOMAIN      467    520       TSP TYPE-1 4.
FT   DOMAIN      522    575       TSP TYPE-1 5.
FT   DOMAIN      881    938       GPS.
FT   DOMAIN     1411   1422       POLY-PRO.
FT   DOMAIN     1425   1430       POLY-PRO.
FT   SITE        231    233       CELL ATTACHMENT SITE (POTENTIAL).
FT   DOMAIN     1365   1584       NECESSARY FOR INTERACTION WITH BAP1.
FT   DOMAIN     1581   1584       INDISPENSABLE FOR INTERACTION WITH BAP1.
FT   CARBOHYD     64     64       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    401    401       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    607    607       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    692    692       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    844    844       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    877    877       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    881    881       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   1584 AA;  173531 MW;  DEA8F28C77874513 CRC64;

  Query Match             5.8%;  Score 280;  DB 1;  Length 1584;
  Best Local Similarity   31.8%;  Pred. No. 3.8e-13;
  Matches   74;  Conservative   38;  Mismatches   93;  Indels   28;  Gaps   10;

Qy        124 CQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAEVEW 183
              |   |   :| | |:   :|    |:   :| |  |:   |    |  |  | | || ||
Db        309 CNREACGPAGRTSSRSQSLRSTDARR---REELGDEL---QQFGFPA-PQTGDPAAE-EW 360

Qy        184 LRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNG 243
                :   |  |     : ||        |  :::|:      :   :|  : : :  |:|
Db        361 --SPWSVCSSTCGEGWQTR--------TRFCVSSSYSTQCSGPLREQRLCNNSAVCPVHG 410

Qy        244 GWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSSLLV---SVDG 300
               |  |: ||:||::||||:: |:|:|     |  ||  |||      :  :  |    :|||
Db        411 AWDEWSPWSLCSSTCGRGFRDRTRTCR--PPQFGGNPCEGPEKQTKFCNIALCPGRAVDG 468

Qy        301 SWSPWSKWSACGLDCT---HWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLC 350
              :|: || ||||   |:      |:|||: |:    || ||||  ::||:|    |
Db        469 NWNEWSSWSACSASCSQGRQQRTRECNGPS--YGGAECQGHWVETRDCFLQQC 519

RESULT 2
TSP2_HUMAN
ID   TSP2_HUMAN     STANDARD;      PRT;  1172 AA.
AC   P35442;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Thrombospondin 2 precursor.
GN   THBS2 OR TSP2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94010892; PubMed=8406456;
RA   Labell T.L., Byers P.H.;
RT   "Sequence and characterization of the complete human thrombospondin 2
RT   cDNA: potential regulatory role for the 3' untranslated region.";
RL   Genomics 17:225-229(1993).
RN   [2]
RP   SEQUENCE OF 560-1172 FROM N.A.
RC   TISSUE=Fibroblast;
RX   MEDLINE=92217961; PubMed=1559694;
RA   Labell T.L., McGookey Milewicz D.J., Disteche C.M., Byers P.H.;
RT   "Thrombospondin II: partial cDNA sequence, chromosome location, and
RT   expression of a second member of the thrombospondin gene family in
RT   humans.";
RL   Genomics 12:421-429(1992).
RN   [3]
RP   THROMBOSPONDIN REPEATS DISULFIDE BONDS.
RX   MEDLINE=21588233; PubMed=11590138;
RA   Misenheimer T.M., Hahr A.J., Harms A.C., Annis D.S., Mosher D.F.;
RT   "Disulfide connectivity of recombinant C-terminal region of human
RT   thrombospondin 2.";
RL   J. Biol. Chem. 276:45882-45887(2001).
CC   -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC       cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC       laminin and type V collagen.
CC   -!- SUBUNIT: Homotrimer; disulfide-linked.
CC   -!- SIMILARITY: Belongs to the thrombospondin family.
CC   -!- SIMILARITY: Contains 1 VWFC domain.
CC   -!- SIMILARITY: Contains 3 EGF-like domains.
CC   -!- SIMILARITY: Contains 3 TSP type-1 domains.
CC   -!- SIMILARITY: Contains 7 TSP type-3 domains.
CC   -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; L12350; AAA03703.1; -.
DR   EMBL; M81339; -; NOT_ANNOTATED_CDS.
DR   PIR; A47379; TSHUP2.
```

```
DR   HSSP; P00740; 1EDM.
DR   Genew; HGNC:11786; THBS2.
DR   MIM; 188061; -.
DR   GO; GO:0008201; F:heparin binding; TAS.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR000884; TSP1.
DR   InterPro; IPR008085; TSP_1.
DR   InterPro; IPR003367; tsp_3.
DR   InterPro; IPR008859; TSPC.
DR   InterPro; IPR003129; TSPN.
DR   InterPro; IPR001007; VWF_C.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00090; tsp_1; 3.
DR   Pfam; PF02412; tsp_3; 13.
DR   Pfam; PF05735; TSPC; 1.
DR   Pfam; PF02210; TSPN; 1.
DR   Pfam; PF00093; vwc; 1.
DR   PRINTS; PR01705; TSP1REPEAT.
DR   SMART; SM00181; EGF; 3.
DR   SMART; SM00209; TSP1; 3.
DR   SMART; SM00210; TSPN; 1.
DR   SMART; SM00214; VWC; 1.
DR   PROSITE; PS00022; EGF_1; FALSE_NEG.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS50026; EGF_3; 2.
DR   PROSITE; PS50092; TSP1; 3.
DR   PROSITE; PS01208; VWFC_1; 1.
DR   PROSITE; PS50184; VWFC_2; 1.
KW   Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat;
KW   EGF-like domain; Signal.
FT   SIGNAL        1     18       POTENTIAL.
FT   CHAIN        19   1172       THROMBOSPONDIN 2.
FT   DOMAIN       19    215       TSP N-TERMINAL.
FT   DOMAIN       19    232       HEPARIN-BINDING (POTENTIAL).
FT   DOMAIN      318    375       VWFC.
FT   DOMAIN      381    431       TSP TYPE-1 1.
FT   DOMAIN      437    492       TSP TYPE-1 2.
FT   DOMAIN      494    549       TSP TYPE-1 3.
FT   DOMAIN      549    589       EGF-LIKE 1.
FT   DOMAIN      590    647       EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      648    692       EGF-LIKE 3.
FT   DOMAIN      725    760       TSP TYPE-3 1.
FT   DOMAIN      761    783       TSP TYPE-3 2.
FT   DOMAIN      784    819       TSP TYPE-3 3.
FT   DOMAIN      820    842       TSP TYPE-3 4.
FT   DOMAIN      843    880       TSP TYPE-3 5.
FT   DOMAIN      881    916       TSP TYPE-3 6.
FT   DOMAIN      917    952       TSP TYPE-3 7.
FT   DOMAIN      953   1172       C-TERMINAL.
FT   SITE        928    930       CELL ATTACHMENT SITE (POTENTIAL).
FT   DISULFID    266    266       INTERCHAIN (PROBABLE).
FT   DISULFID    270    270       INTERCHAIN (PROBABLE).
FT   DISULFID    393    425       BY SIMILARITY.
FT   DISULFID    397    430       BY SIMILARITY.
FT   DISULFID    408    415       BY SIMILARITY.
FT   DISULFID    449    486       BY SIMILARITY.
FT   DISULFID    453    491       BY SIMILARITY.
FT   DISULFID    464    476       BY SIMILARITY.
FT   DISULFID    506    543       BY SIMILARITY.
FT   DISULFID    510    548       BY SIMILARITY.
FT   DISULFID    521    533       BY SIMILARITY.
FT   DISULFID    553    564       BY SIMILARITY.
FT   DISULFID    558    574       BY SIMILARITY.
FT   DISULFID    577    588       BY SIMILARITY.
FT   DISULFID    594    610       BY SIMILARITY.
FT   DISULFID    601    619       BY SIMILARITY.
FT   DISULFID    622    646       BY SIMILARITY.
FT   DISULFID    652    665       BY SIMILARITY.
FT   DISULFID    659    678       BY SIMILARITY.
FT   DISULFID    680    691       BY SIMILARITY.
FT   DISULFID    707    715
FT   DISULFID    720    740
FT   DISULFID    756    776
FT   DISULFID    779    799
FT   DISULFID    815    835
FT   DISULFID    838    858
FT   DISULFID    876    896
FT   DISULFID    912    932
FT   DISULFID    948   1169
FT   CARBOHYD    151    151       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    316    316       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    330    330       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    457    457       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    584    584       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    710    710       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1069   1069       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   1172 AA;  129955 MW;  2AC7BB230E44C6F5 CRC64;

  Query Match             5.7%;  Score 273.5;  DB 1;  Length 1172;
  Best Local Similarity   28.9%;  Pred. No. 7.9e-13;
  Matches   74;  Conservative   27;  Mismatches  110;  Indels   45;  Gaps    8;

Qy        209 RQARLADTANYTCVAKNIVARRRSASAA-VIVYVNGGWSTWTEWSVCSASCGRGWQKRSR 267
              :: |  |  : ||:  :|  |  | |     :  :|||| |: || || :|| |   | |
Db        403 QRGRSCDVTSNTCLGPSIQTRACSLSKCDTRIRQDGGWSHWSPWSSCSVTCGVGNITRIR 462

Qy        268 SCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDGSWSPWSKWSACGLDCT---HWRSRECS 324
               | :| |  ||  |:|     :        :|| ||||| |||| : |      |:| |:
Db        463 LCNSPVPQMGGKNCKGSGRETKACQGAPCPIDGRWSPWSPWSACTVTCAGGIRERTRVCN 522

Qy        325 DPAPRNGGEECQGTDLDTRNCTSDLCVHSASGPEDVALYVGLIAVAVCLVLLLLVLILVY 384
               | |: ||: | |   : : |    |      | |  |       | |
Db        523 SPEPQYGGKACVGDVQERQMCNKRSC------PVDGCLSNPCFPGAQC------------ 564

Qy        385 CRKKEGLDSDVADSSILTSGFQPVSI--KPSKADNPHLLTIQPDLSTTT--------TYQ 434
                      |   | |  : || ||      :  ::     : ||:  :|        |
Db        565 --------SSFPDGS-WSCGFCPVGFLGNGTHCEDLDECALVPDICFSTSKVPRCVNTQP 615

Qy        435 GSLC----PRQDGPSP 446
              |  |    ||  |  |
Db        616 GFHCLPCPPRYRGNQP 631


RESULT 3
TSP2_MOUSE
ID   TSP2_MOUSE     STANDARD;      PRT;  1172 AA.
AC   Q03350;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Thrombospondin 2 precursor.
GN   THBS2 OR TSP2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92147683; PubMed=1371115;
RA   Laherty C.D., O'Rourke K., Wolf F.W., Katz R., Seldin M.F.,
RA   Dixit V.M.;
RT   "Characterization of mouse thrombospondin 2 sequence and expression
RT   during cell growth and development.";
RL   J. Biol. Chem. 267:3274-3281(1992).
RN   [2]
RP   SEQUENCE OF 1-873 FROM N.A.
RX   MEDLINE=91302287; PubMed=1712771;
RA   Bornstein P., O'Rourke K., Wikstrom K., Wolf F.W., Katz R., Li P.,
RA   Dixit V.M.;
RT   "A second, expressed thrombospondin gene (Thbs2) exists in the mouse
RT   genome.";
RL   J. Biol. Chem. 266:12821-12824(1991).
CC   -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC       cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
```

```
CC       laminin and type V collagen.
CC   -!- SUBUNIT: Homotrimer; disulfide-linked.
CC   -!- SIMILARITY: Belongs to the thrombospondin family.
CC   -!- SIMILARITY: Contains 1 VWFC domain.
CC   -!- SIMILARITY: Contains 3 EGF-like domains.
CC   -!- SIMILARITY: Contains 3 TSP type-1 domains.
CC   -!- SIMILARITY: Contains 7 TSP type-3 domains.
CC   -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; L07803; AAA53064.1; -.
DR   EMBL; M64866; AAA40432.1; -.
DR   PIR; A42587; A42587.
DR   HSSP; P00740; 1EDM.
DR   MGD; MGI:98738; Thbs2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR000884; TSP1.
DR   InterPro; IPR008085; TSP_1.
DR   InterPro; IPR003367; tsp_3.
DR   InterPro; IPR008859; TSPC.
DR   InterPro; IPR003129; TSPN.
DR   InterPro; IPR001007; VWF_C.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00090; tsp_1; 3.
DR   Pfam; PF02412; tsp_3; 13.
DR   Pfam; PF05735; TSPC; 1.
DR   Pfam; PF02210; TSPN; 1.
DR   Pfam; PF00093; vwc; 1.
DR   PRINTS; PR01705; TSP1REPEAT.
DR   SMART; SM00181; EGF; 3.
DR   SMART; SM00209; TSP1; 3.
DR   SMART; SM00210; TSPN; 1.
DR   SMART; SM00214; VWC; 1.
DR   PROSITE; PS00022; EGF_1; FALSE_NEG.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS50026; EGF_3; 2.
DR   PROSITE; PS50092; TSP1; 3.
DR   PROSITE; PS01208; VWFC_1; 1.
DR   PROSITE; PS50184; VWFC_2; 1.
KW   Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat;
KW   EGF-like domain; Signal.
FT   SIGNAL        1     18       POTENTIAL.
FT   CHAIN        19   1172       THROMBOSPONDIN 2.
FT   DOMAIN       19    215       TSP N-TERMINAL.
FT   DOMAIN       19    232       HEPARIN-BINDING (POTENTIAL).
FT   DOMAIN      318    375       VWFC.
FT   DOMAIN      381    431       TSP TYPE-1 1.
FT   DOMAIN      437    492       TSP TYPE-1 2.
FT   DOMAIN      494    549       TSP TYPE-1 3.
FT   DOMAIN      549    589       EGF-LIKE 1.
FT   DOMAIN      590    647       EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      648    692       EGF-LIKE 3.
FT   DOMAIN      725    760       TSP TYPE-3 1.
FT   DOMAIN      761    783       TSP TYPE-3 2.
FT   DOMAIN      784    819       TSP TYPE-3 3.
FT   DOMAIN      820    842       TSP TYPE-3 4.
FT   DOMAIN      843    880       TSP TYPE-3 5.
FT   DOMAIN      881    916       TSP TYPE-3 6.
FT   DOMAIN      917    952       TSP TYPE-3 7.
FT   DOMAIN      953   1172       C-TERMINAL.
FT   SITE        928    930       CELL ATTACHMENT SITE (POTENTIAL).
FT   DISULFID    266    266       INTERCHAIN (PROBABLE).
FT   DISULFID    270    270       INTERCHAIN (PROBABLE).
FT   DISULFID    393    425       BY SIMILARITY.
FT   DISULFID    397    430       BY SIMILARITY.
FT   DISULFID    408    415       BY SIMILARITY.
FT   DISULFID    449    486       BY SIMILARITY.
FT   DISULFID    453    491       BY SIMILARITY.
FT   DISULFID    464    476       BY SIMILARITY.
FT   DISULFID    506    543       BY SIMILARITY.
FT   DISULFID    510    548       BY SIMILARITY.
FT   DISULFID    521    533       BY SIMILARITY.
FT   DISULFID    553    564       BY SIMILARITY.
FT   DISULFID    558    574       BY SIMILARITY.
FT   DISULFID    577    588       BY SIMILARITY.
FT   DISULFID    594    610       BY SIMILARITY.
FT   DISULFID    601    619       BY SIMILARITY.
FT   DISULFID    622    646       BY SIMILARITY.
FT   DISULFID    652    665       BY SIMILARITY.
FT   DISULFID    659    678       BY SIMILARITY.
FT   DISULFID    680    691       BY SIMILARITY.
FT   DISULFID    707    715       BY SIMILARITY.
FT   DISULFID    720    740       BY SIMILARITY.
FT   DISULFID    756    776       BY SIMILARITY.
FT   DISULFID    779    799       BY SIMILARITY.
FT   DISULFID    815    835       BY SIMILARITY.
FT   DISULFID    838    858       BY SIMILARITY.
FT   DISULFID    876    896       BY SIMILARITY.
FT   DISULFID    912    932       BY SIMILARITY.
FT   DISULFID    948   1169       BY SIMILARITY.
FT   CARBOHYD    151    151       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    316    316       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    330    330       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    457    457       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    584    584       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    710    710       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1069   1069       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   1172 AA;  129911 MW;  7CE8E4E8599822AB CRC64;

  Query Match             5.7%;  Score 273.5;  DB 1;  Length 1172;
  Best Local Similarity   36.1%;  Pred. No. 7.9e-13;
  Matches   57;  Conservative   22;  Mismatches   70;  Indels    9;  Gaps    4;

Qy        209 RQARLADTANYTCVAKNIVARRRS-ASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSR 267
              :: |  |  : ||: :|  |  |        :  ||||| |: || || :||  |   | |
Db        403 QRGRSCDVTSNTCLGPSIQTRTCSLGKCDTRIRQNGGWSHWSPWSSCSVTCGVGNVTRIR 462

Qy        268 SCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDGSWSPWSKWSACGLDCT---HWRSRECS 324
              | :|  |  ||  |:|        :         :|| ||||| ||||  : |       ||| |:
Db        463 LCNSPVPQMGGKNCKGSGRETKPCQRDPCPIDGRWSPWSPWSACTVTCAGGIRERSRVCN 522

Qy        325 DPAPRNGGEECQG--TD---LDTRNCTSDLCVHSASGP 357
              | |: ||::|  |  |:    : |:|  | |: :   |
Db        523 SPEPQYGGKDCVGDVTEHQMCNKRSCPIDGCLSNPCFP 560


RESULT 4
SM5A_MOUSE
ID   SM5A_MOUSE     STANDARD;      PRT;  1077 AA.
AC   Q62217;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Semaphorin 5A precursor (Semaphorin F) (Sema F).
GN   SEMA5A OR SEMAF OR SEMF.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NMRI;
RX   MEDLINE=96414430; PubMed=8817451;
RA   Adams R.H., Betz H., Pueschel A.W.;
RT   "A novel class of murine semaphorins with homology to thrombospondin
RT   is differentially expressed during early embryogenesis.";
RL   Mech. Dev. 57:33-45(1996).
```

```
CC   -!- FUNCTION: May act as positive axonal guidance cues.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- TISSUE SPECIFICITY: IN ADULT, DETECTED IN LIVER, BRAIN, KIDNEY,
CC       HEART, LUNG AND SPLEEN.
CC   -!- DEVELOPMENTAL STAGE: DIFFERENTIALLY EXPRESSED IN EMBRYONIC AND
CC       ADULT TISSUES. ITS ABUNDANCE DECREASES FROM E10 TO BITH.
CC   -!- SIMILARITY: Belongs to the semaphorin family.
CC   -!- SIMILARITY: Contains 1 Sema domain.
CC   -!- SIMILARITY: Contains 7 TSP type-1 domains.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X97817; CAA66397.1; -.
DR   MGD; MGI:107556; Sema5a.
DR   GO; GO:0016021; C:integral to membrane; IDA.
DR   GO; GO:0008046; F:axon guidance receptor activity; IDA.
DR   GO; GO:0007411; P:axon guidance; IMP.
DR   InterPro; IPR003659; Plexin-like.
DR   InterPro; IPR002165; Plexin_repeat.
DR   InterPro; IPR001627; Sema.
DR   InterPro; IPR000884; TSP1.
DR   InterPro; IPR008085; TSP_1.
DR   Pfam; PF01437; PSI; 1.
DR   Pfam; PF01403; Sema; 1.
DR   Pfam; PF00090; tsp_1; 5.
DR   PRINTS; PR01705; TSP1REPEAT.
DR   SMART; SM00423; PSI; 1.
DR   SMART; SM00630; Sema; 1.
DR   SMART; SM00209; TSP1; 6.
DR   PROSITE; PS50092; TSP1; 6.
KW   Signal; Transmembrane; Repeat; Multigene family; Neurogenesis;
KW   Developmental protein; Glycoprotein.
FT   SIGNAL        1     21       POTENTIAL.
FT   CHAIN        22   1077       SEMAPHORIN 5A.
FT   DOMAIN       22    971       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    972    992       POTENTIAL.
FT   DOMAIN      993   1077       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      226    507       SEMA.
FT   DOMAIN      540    593       TSP TYPE-1 1.
FT   DOMAIN      595    651       TSP TYPE-1 2.
FT   DOMAIN      653    702       TSP TYPE-1 3.
FT   DOMAIN      707    765       TSP TYPE-1 4.
FT   DOMAIN      784    839       TSP TYPE-1 5.
FT   DOMAIN      841    896       TSP TYPE-1 6.
FT   DOMAIN      897    944       TSP TYPE-1 7.
FT   CARBOHYD    147    147       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    168    168       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    227    227       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    277    277       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    323    323       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    367    367       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    536    536       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    591    591       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    717    717       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    933    933       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   1077 AA;  120826 MW;  EDAB0DDDA42789FF CRC64;

  Query Match             5.7%;  Score 272.5;  DB 1;  Length 1077;
  Best Local Similarity   43.2%;  Pred. No. 8.4e-13;
  Matches   51;  Conservative   11;  Mismatches   53;  Indels    3;  Gaps    1;

Qy        241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDG 300
              ||| || || || ||  | || :  | | | ||  |  ||  |  | ::  :  : |   |||
Db        783 VNGAWSAWTSWSQCSRDCSRGIRNRKRVCNNPEPKFGGMPCLGPSLEFQECNILPCPVDG 842

Qy        301 SWSPWSKWSACGLDC---THWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCVHSAS 355
              || || || |  |     : |:| ||:|||  ||: | |   :    | :  |  | |

Db        843 VWSCWSSWSKCSATCGGGHYMRTRSCSNPAPAYGGDICLGLHTEEALCNTQTCPESWS 900


RESULT 5
SM5A_HUMAN
ID   SM5A_HUMAN     STANDARD;      PRT;  1074 AA.
AC   Q13591; O60408;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Semaphorin 5A precursor (Semaphorin F) (Sema F).
GN   SEMA5A OR SEMAF.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98125554; PubMed=9464278;
RA   Simmons A.D., Puschel A.W., McPherson J.D., Overhauser J., Lovett M.;
RT   "Molecular cloning and mapping of human semaphorin F from the Cri-du-
RT   chat candidate interval.";
RL   Biochem. Biophys. Res. Commun. 242:685-691(1998).
RN   [2]
RP   SEQUENCE OF 1-494 FROM N.A.
RA   Kalicki J., Harmon G.;
RL   Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: May act as positive axonal guidance cues.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- SIMILARITY: Belongs to the semaphorin family.
CC   -!- SIMILARITY: Contains 1 Sema domain.
CC   -!- SIMILARITY: Contains 7 TSP type-1 domains.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U52840; AAC09473.1; -.
DR   EMBL; AC004615; AAC14668.1; -.
DR   PIR; JC5928; JC5928.
DR   Genew; HGNC:10736; SEMA5A.
DR   GO; GO:0007155; P:cell adhesion; TAS.
DR   GO; GO:0007267; P:cell-cell signaling; TAS.
DR   GO; GO:0007399; P:neurogenesis; TAS.
DR   InterPro; IPR003659; Plexin-like.
DR   InterPro; IPR002165; Plexin_repeat.
DR   InterPro; IPR001627; Sema.
DR   InterPro; IPR000884; TSP1.
DR   InterPro; IPR008085; TSP_1.
DR   Pfam; PF01437; PSI; 1.
DR   Pfam; PF01403; Sema; 1.
DR   Pfam; PF00090; tsp_1; 6.
DR   PRINTS; PR01705; TSP1REPEAT.
DR   SMART; SM00423; PSI; 1.
DR   SMART; SM00630; Sema; 1.
DR   SMART; SM00209; TSP1; 6.
DR   PROSITE; PS50092; TSP1; 6.
KW   Signal; Transmembrane; Repeat; Multigene family; Neurogenesis;
KW   Developmental protein; Glycoprotein.
FT   SIGNAL        1     22       POTENTIAL.
FT   CHAIN        23   1074       SEMAPHORIN 5A.
FT   DOMAIN       23    968       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    969    989       POTENTIAL.
FT   DOMAIN      990   1074       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      226    507       SEMA.
FT   DOMAIN      540    593       TSP TYPE-1 1.
FT   DOMAIN      595    651       TSP TYPE-1 2.
FT   DOMAIN      653    702       TSP TYPE-1 3.
FT   DOMAIN      707    765       TSP TYPE-1 4.
```

```
FT   DOMAIN      784    839       TSP TYPE-1 5.
FT   DOMAIN      841    896       TSP TYPE-1 6.
FT   DOMAIN      897    944       TSP TYPE-1 7.
FT   CARBOHYD    142    142       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    168    168       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    227    227       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    277    277       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    323    323       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    367    367       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    437    437       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    536    536       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    591    591       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    717    717       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    933    933       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT     56     56       A -> V (IN REF. 2).
FT   CONFLICT    149    149       A -> T (IN REF. 2).
FT   CONFLICT    382    382       V -> M (IN REF. 2).
FT   CONFLICT    494    494       S -> R (IN REF. 2).
SQ   SEQUENCE   1074 AA;  120570 MW;  EE3DB763CBE29407 CRC64;

  Query Match             5.7%;  Score 270.5;  DB 1;  Length 1074;
  Best Local Similarity   42.4%;  Pred. No. 1.2e-12;
  Matches   50;  Conservative   13;  Mismatches   52;  Indels    3;  Gaps    1;

Qy        241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDG 300
              ||| || || || ||  | || : | | | || |  ||  | | ::  :  ::|   |||
Db        783 VNGAWSAWTSWSQCSRDCSRGIRNRKRVCNNPEPKYGGMPCLGPSLEYQECNTLPCPVDG 842

Qy        301 SWSPWSKWSACGLDC---THWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCVHSAS 355
               || || |: |   |     : |:| ||:|||  ||: | |    :  | :  |  | |
Db        843 VWSCWSPWTKCSATCGGGHYMRTRSCSNPAPAYGGDICLGLHTEEALCNTQPCPESWS 900


RESULT 6
TSP2_BOVIN
ID   TSP2_BOVIN     STANDARD;      PRT;  1170 AA.
AC   Q95116; Q28180;
DT   01-NOV-1997 (Rel. 35, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Thrombospondin 2 precursor (Corticotropin-induced secreted protein)
DE   (CISP).
GN   THBS2 OR TSP2 OR TSP-2.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Danik M., Chinn A., Lafeuillade M., Keramidas M., Aguesse-Germon S.,
RA   Penhoat A., Chen H., Mosher D., Chambaz E.M., Feige J.J.;
RL   Submitted (MAR-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 1-522 FROM N.A.
RX   MEDLINE=96331130; PubMed=8698834;
RA   Lafeuillade B., Pellerin S., Keramidas M., Danik M., Chambaz E.M.,
RA   Feige J.J.;
RT   "Opposite regulation of thrombospondin-1 and corticotropin-induced
RT   secreted protein/thrombospondin-2 expression by adrenocorticotropic
RT   hormone in adrenocortical cells.";
RL   J. Cell. Physiol. 167:164-172(1996).
RN   [3]
RP   SEQUENCE OF 318-831 FROM N.A.
RC   TISSUE=Aortic endothelium;
RA   Zafar R.S., Moll Y.D., Womack J.F., Walz D.A.;
RT   "Cloning and sequencing of bovine thrombospondin stimulatory effect of
RT   TGF-beta.";
RL   Submitted (MAY-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC       cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC       laminin and type V collagen.
CC   -!- SUBUNIT: Homotrimer; disulfide-linked.
CC   -!- SIMILARITY: Belongs to the thrombospondin family.
CC   -!- SIMILARITY: Contains 1 VWFC domain.
CC   -!- SIMILARITY: Contains 3 EGF-like domains.
CC   -!- SIMILARITY: Contains 3 TSP type-1 domains.
CC   -!- SIMILARITY: Contains 7 TSP type-3 domains.
CC   -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X96540; CAA65385.1; -.
DR   EMBL; X87620; CAA60952.1; -.
DR   HSSP; P00740; 1EDM.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR000884; TSP1.
DR   InterPro; IPR008085; TSP_1.
DR   InterPro; IPR003367; tsp_3.
DR   InterPro; IPR008859; TSPC.
DR   InterPro; IPR003129; TSPN.
DR   InterPro; IPR001007; VWF_C.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF00090; tsp_1; 3.
DR   Pfam; PF02412; tsp_3; 13.
DR   Pfam; PF05735; TSPC; 1.
DR   Pfam; PF02210; TSPN; 1.
DR   Pfam; PF00093; vwc; 1.
DR   PRINTS; PR01705; TSP1REPEAT.
DR   SMART; SM00181; EGF; 3.
DR   SMART; SM00209; TSP1; 3.
DR   SMART; SM00210; TSPN; 1.
DR   SMART; SM00214; VWC; 1.
DR   PROSITE; PS00022; EGF_1; FALSE_NEG.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS50026; EGF_3; 2.
DR   PROSITE; PS50092; TSP1; 3.
DR   PROSITE; PS01208; VWFC_1; 1.
DR   PROSITE; PS50184; VWFC_2; 1.
KW   Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat;
KW   EGF-like domain; Signal.
FT   SIGNAL        1     18       POTENTIAL.
FT   CHAIN        19   1170       THROMBOSPONDIN 2.
FT   DOMAIN       19    215       TSP N-TERMINAL.
FT   DOMAIN       19    232       HEPARIN-BINDING (POTENTIAL).
FT   DOMAIN      318    375       VWFC.
FT   DOMAIN      379    429       TSP TYPE-1 1.
FT   DOMAIN      435    490       TSP TYPE-1 2.
FT   DOMAIN      492    547       TSP TYPE-1 3.
FT   DOMAIN      547    587       EGF-LIKE 1.
FT   DOMAIN      588    645       EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      646    690       EGF-LIKE 3.
FT   DOMAIN      723    758       TSP TYPE-3 1.
FT   DOMAIN      759    781       TSP TYPE-3 2.
FT   DOMAIN      782    817       TSP TYPE-3 3.
FT   DOMAIN      818    840       TSP TYPE-3 4.
FT   DOMAIN      841    878       TSP TYPE-3 5.
FT   DOMAIN      879    914       TSP TYPE-3 6.
FT   DOMAIN      915    950       TSP TYPE-3 7.
FT   DOMAIN      951   1170       C-TERMINAL.
FT   SITE        926    928       CELL ATTACHMENT SITE (POTENTIAL).
FT   DISULFID    266    266       INTERCHAIN (PROBABLE).
FT   DISULFID    270    270       INTERCHAIN (PROBABLE).
FT   DISULFID    391    423       BY SIMILARITY.
FT   DISULFID    395    428       BY SIMILARITY.
FT   DISULFID    406    413       BY SIMILARITY.
FT   DISULFID    447    484       BY SIMILARITY.
FT   DISULFID    451    489       BY SIMILARITY.
```

```
FT   DISULFID    462    474       BY SIMILARITY.
FT   DISULFID    504    541       BY SIMILARITY.
FT   DISULFID    508    546       BY SIMILARITY.
FT   DISULFID    519    531       BY SIMILARITY.
FT   DISULFID    551    562       BY SIMILARITY.
FT   DISULFID    556    572       BY SIMILARITY.
FT   DISULFID    575    586       BY SIMILARITY.
FT   DISULFID    592    608       BY SIMILARITY.
FT   DISULFID    599    617       BY SIMILARITY.
FT   DISULFID    620    644       BY SIMILARITY.
FT   DISULFID    650    663       BY SIMILARITY.
FT   DISULFID    657    676       BY SIMILARITY.
FT   DISULFID    678    689       BY SIMILARITY.
FT   DISULFID    705    713       BY SIMILARITY.
FT   DISULFID    718    738       BY SIMILARITY.
FT   DISULFID    754    774       BY SIMILARITY.
FT   DISULFID    777    797       BY SIMILARITY.
FT   DISULFID    813    833       BY SIMILARITY.
FT   DISULFID    836    856       BY SIMILARITY.
FT   DISULFID    874    894       BY SIMILARITY.
FT   DISULFID    910    930       BY SIMILARITY.
FT   DISULFID    946   1167       BY SIMILARITY.
FT   CARBOHYD    151    151       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    316    316       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    330    330       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    455    455       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    582    582       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    708    708       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    936    936       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1067   1067       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    535    535       A -> V (IN REF. 3).
FT   CONFLICT    748    748       S -> T (IN REF. 3).
SQ   SEQUENCE   1170 AA;  129862 MW;  9CF1FBF55B89A051 CRC64;

  Query Match             5.6%;  Score 269;  DB 1;  Length 1170;
  Best Local Similarity   35.6%;  Pred. No. 1.8e-12;
  Matches   52;  Conservative   21;  Mismatches   69;  Indels    4;  Gaps    2;

Qy        209 RQARLADTANYTCVAKNIVARRRSASAA-VIVYVNGGWSTWTEWSVCSASCGRGWQKRSR 267
              :: |  |  : ||:  :|  |  |       :  :|||| |: || || :|| |   | |
Db        401 QRGRSCDVTSNTCLGPSIQTRACSLGRCDHRIRQDGGWSHWSPWSSCSVTCGVGNVTRIR 460

Qy        268 SCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDGSWSPWSKWSACGLDCT---HWRSRECS 324
               | :| |  ||  |:|     :        ||| ||||| |||| : |      |:| |:
Db        461 LCNSPVPQMGGRSCKGSGRETKACQGPPCPVDGRWSPWSPWSACTVTCAGGIRERTRVCN 520

Qy        325 DPAPRNGGEECQGTDLDTRNCTSDLC 350
               | |::||::| |   : : |    |
Db        521 SPEPQHGGKDCVGGAKEQQMCNRKSC 546


RESULT 7
BAI3_HUMAN
ID   BAI3_HUMAN     STANDARD;      PRT;  1522 AA.
AC   O60242; O60297;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Brain-specific angiogenesis inhibitor 3 precursor.
GN   BAI3 OR KIAA0550.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fetal brain;
RX   MEDLINE=98194217; PubMed=9533023;
RA   Shiratsuchi T., Nishimori H., Ichise H., Nakamura Y., Tokino T.;
RT   "Cloning and characterization of BAI2 and BAI3, novel genes homologous
RT   to brain-specific angiogenesis inhibitor 1 (BAI1).";
RL   Cytogenet. Cell Genet. 79:103-108(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=98290545; PubMed=9628581;
RA   Nagase T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H.,
RA   Nomura N., Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes. IX.
RT   The complete sequences of 100 new cDNA clones from brain which can
RT   code for large proteins in vitro.";
RL   DNA Res. 5:31-39(1998).
RN   [3]
RP   SEQUENCE FROM N.A., AND REVISIONS TO 643-665 AND C-TERMINUS.
RX   MEDLINE=22158633; PubMed=12168954;
RA   Nakajima D., Okazaki N., Yamakawa H., Kikuno R., Ohara O., Nagase T.;
RT   "Construction of expression-ready cDNA clones for KIAA genes: manual
RT   curation of 330 KIAA cDNA clones.";
RL   DNA Res. 9:99-106(2002).
CC   -!- FUNCTION: MIGHT BE INVOLVED IN ANGIOGENESIS INHIBITION AND
CC       SUPPRESSION OF GLIOBLASTOMA.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: STRONGLY EXPRESSED IN BRAIN. ALSO DETECTED IN
CC       HEART. REDUCED EXPRESSION IS OBSERVED IN SOME GLIOBLASTOMA CELL
CC       LINES.
CC   -!- SIMILARITY: Belongs to family 2 of G-protein coupled receptors.
CC   -!- SIMILARITY: Contains 1 CUB domain.
CC   -!- SIMILARITY: Contains 4 TSP type-1 domains.
CC   -!- SIMILARITY: Contains 1 GPS domain.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AB005299; BAA25363.1; -.
DR   EMBL; AB011122; BAA25476.2; ALT_INIT.
DR   PIR; T00028; T00028.
DR   Genew; HGNC:945; BAI3.
DR   MIM; 602684; -.
DR   InterPro; IPR000859; CUB.
DR   InterPro; IPR000832; GPCR_secretin.
DR   InterPro; IPR001879; hormn_receptor.
DR   InterPro; IPR000203; PKD_cys_rich.
DR   InterPro; IPR000884; TSP1.
DR   Pfam; PF00002; 7tm_2; 1.
DR   Pfam; PF01825; GPS; 1.
DR   Pfam; PF02793; HRM; 1.
DR   Pfam; PF00090; tsp_1; 4.
DR   SMART; SM00303; GPS; 1.
DR   SMART; SM00008; HormR; 1.
DR   SMART; SM00209; TSP1; 4.
DR   PROSITE; PS01180; CUB; 1.
DR   PROSITE; PS50221; GPS; 1.
DR   PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; FALSE_NEG.
DR   PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; FALSE_NEG.
DR   PROSITE; PS50227; G_PROTEIN_RECEP_F2_3; 1.
DR   PROSITE; PS50261; G_PROTEIN_RECEP_F2_4; 1.
DR   PROSITE; PS50092; TSP1; 4.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW   Repeat.
FT   SIGNAL         1     24       POTENTIAL.
FT   CHAIN         25   1522       BRAIN-SPECIFIC ANGIOGENESIS INHIBITOR 3.
FT   DOMAIN        25    880       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     881    901       1 (POTENTIAL).
FT   DOMAIN       902    910       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     911    931       2 (POTENTIAL).
FT   DOMAIN       932    939       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     940    960       3 (POTENTIAL).
FT   DOMAIN       961    981       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     982   1002       4 (POTENTIAL).
FT   DOMAIN      1003   1023       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    1024   1044       5 (POTENTIAL).
```

```
FT   DOMAIN      1045   1098       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    1099   1119       6 (POTENTIAL).
FT   DOMAIN      1120   1125       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    1126   1146       7 (POTENTIAL).
FT   DOMAIN      1147   1522       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN        30    159       CUB.
FT   DOMAIN       291    343       TSP TYPE-1 1.
FT   DOMAIN       345    398       TSP TYPE-1 2.
FT   DOMAIN       400    453       TSP TYPE-1 3.
FT   DOMAIN       455    508       TSP TYPE-1 4.
FT   DOMAIN       816    868       GPS.
FT   DOMAIN       942    945       POLY-THR.
FT   DOMAIN      1173   1176       POLY-SER.
FT   CARBOHYD      51     51       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      54     54       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      82     82       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     105    105       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     241    241       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     337    337       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     418    418       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     540    540       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     625    625       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     779    779       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     812    812       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     828    828       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     937    937       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   1522 AA;  171490 MW;  D22D0A5D4BB62502 CRC64;

  Query Match             5.6%;  Score 267.5;  DB 1;  Length 1522;
  Best Local Similarity   37.7%;  Pred. No. 3.3e-12;
  Matches   55;  Conservative   22;  Mismatches   54;  Indels   15;  Gaps    5;

Qy        220 TCVA------KNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPA 273
              |||:         :   |  :   :  |:| |  |: ||:|| :|||| : |:||||
Db        317 TCVSPYGTHCSGPLRESRVCNNTALCPVHGVWEEWSPWSLCSFTCGRGQRTRTRSCT--P 374

Qy        274 PLNGGAFCEGQNVHDRTVSSLLVSVDGSWSPWSKWSACGLDC---THWRSRECSDPAPRN 330
              |  ||  |||   | :  :  |  ||| |  || || | : |   |  |||:|:  |  :
Db        375 PQYGGRPCEGPETHHKPCNIALCPVDGQWQEWSSWSQCSVTCSNGTQQRSRQCT--AAAH 432

Qy        331 GGEECQGTDLDTRNCTSDLCVHSASG 356
              || ||:|   ::| | :  |  :|:|
Db        433 GGSECRGPWAESRECYNPEC--TANG 456


RESULT 8
TSP1_XENLA
ID   TSP1_XENLA     STANDARD;      PRT;  1173 AA.
AC   P35448;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Thrombospondin 1 precursor.
GN   THBS1 OR TSP1.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Urry L.A., Ramos J., Duquette M., Desimone D.W., Lawler J.;
RL   Submitted (XXX-1993) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC       cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC       laminin, type V collagen and integrins alpha-V/beta-1, alpha-
CC       V/beta-3 and alpha-IIb/beta-3 (By similarity).
CC   -!- SUBUNIT: Homotrimer; disulfide-linked.
CC   -!- SIMILARITY: Belongs to the thrombospondin family.
CC   -!- SIMILARITY: Contains 1 VWFC domain.
CC   -!- SIMILARITY: Contains 3 EGF-like domains.
CC   -!- SIMILARITY: Contains 3 TSP type-1 domains.
CC   -!- SIMILARITY: Contains 7 TSP type-3 domains.
CC   -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; L04278; -; NOT_ANNOTATED_CDS.
DR   HSSP; P00740; 1EDM.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR000884; TSP1.
DR   InterPro; IPR008085; TSP_1.
DR   InterPro; IPR003367; tsp_3.
DR   InterPro; IPR008859; TSPC.
DR   InterPro; IPR003129; TSPN.
DR   InterPro; IPR001007; VWF_C.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00090; tsp_1; 3.
DR   Pfam; PF02412; tsp_3; 13.
DR   Pfam; PF05735; TSPC; 1.
DR   Pfam; PF02210; TSPN; 1.
DR   Pfam; PF00093; vwc; 1.
DR   PRINTS; PR01705; TSP1REPEAT.
DR   SMART; SM00181; EGF; 2.
DR   SMART; SM00209; TSP1; 3.
DR   SMART; SM00210; TSPN; 1.
DR   SMART; SM00214; VWC; 1.
DR   PROSITE; PS00022; EGF_1; FALSE_NEG.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS50026; EGF_3; 2.
DR   PROSITE; PS50092; TSP1; 3.
DR   PROSITE; PS01208; VWFC_1; 1.
DR   PROSITE; PS50184; VWFC_2; 1.
KW   Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat;
KW   EGF-like domain; Signal.
FT   SIGNAL         1     22       POTENTIAL.
FT   CHAIN         23   1173       THROMBOSPONDIN 1.
FT   DOMAIN        23    235       HEPARIN-BINDING (POTENTIAL).
FT   DOMAIN        23    224       TSP N-TERMINAL.
FT   DOMAIN       319    376       VWFC.
FT   DOMAIN       382    432       TSP TYPE-1 1.
FT   DOMAIN       438    493       TSP TYPE-1 2.
FT   DOMAIN       495    550       TSP TYPE-1 3.
FT   DOMAIN       550    590       EGF-LIKE 1.
FT   DOMAIN       591    648       EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN       649    693       EGF-LIKE 3.
FT   DOMAIN       726    761       TSP TYPE-3 1.
FT   DOMAIN       762    784       TSP TYPE-3 2.
FT   DOMAIN       785    820       TSP TYPE-3 3.
FT   DOMAIN       821    843       TSP TYPE-3 4.
FT   DOMAIN       844    881       TSP TYPE-3 5.
FT   DOMAIN       882    917       TSP TYPE-3 6.
FT   DOMAIN       918    953       TSP TYPE-3 7.
FT   DOMAIN       954   1173       C-TERMINAL.
FT   SITE         929    931       CELL ATTACHMENT SITE (POTENTIAL).
FT   DISULFID     394    426       BY SIMILARITY.
FT   DISULFID     398    431       BY SIMILARITY.
FT   DISULFID     409    416       BY SIMILARITY.
FT   DISULFID     450    487       BY SIMILARITY.
FT   DISULFID     454    492       BY SIMILARITY.
FT   DISULFID     465    477       BY SIMILARITY.
FT   DISULFID     507    544       BY SIMILARITY.
FT   DISULFID     511    549       BY SIMILARITY.
FT   DISULFID     522    534       BY SIMILARITY.
FT   DISULFID     554    565       BY SIMILARITY.
FT   DISULFID     559    575       BY SIMILARITY.
FT   DISULFID     578    589       BY SIMILARITY.
FT   DISULFID     595    611       BY SIMILARITY.
```

```
FT   DISULFID    602    620       BY SIMILARITY.
FT   DISULFID    623    647       BY SIMILARITY.
FT   DISULFID    653    666       BY SIMILARITY.
FT   DISULFID    660    679       BY SIMILARITY.
FT   DISULFID    681    692       BY SIMILARITY.
FT   DISULFID    708    716       BY SIMILARITY.
FT   DISULFID    721    741       BY SIMILARITY.
FT   DISULFID    757    777       BY SIMILARITY.
FT   DISULFID    780    800       BY SIMILARITY.
FT   DISULFID    816    836       BY SIMILARITY.
FT   DISULFID    839    859       BY SIMILARITY.
FT   DISULFID    877    897       BY SIMILARITY.
FT   DISULFID    913    933       BY SIMILARITY.
FT   DISULFID    949   1170       BY SIMILARITY.
FT   CARBOHYD    155    155       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    158    158       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    250    250       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    363    363       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    705    705       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    711    711       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1070   1070       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   1173 AA;  130019 MW;  A9F036D6516C0F24 CRC64;

  Query Match              5.6%;  Score 266.5;  DB 1;  Length 1173;
  Best Local Similarity   22.5%;  Pred. No. 2.7e-12;
  Matches   86;  Conservative   54;  Mismatches  147;  Indels   95;  Gaps   14;

Qy         11 LLGIVLAAWLRGSG----AQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVC- 65
              : |  | | || |             |: ||| |::| :  :::         | |  :  ||
Db        221 VFGTTLEAILRNKGCLSMTNSVITLDNPVNGSSPAIRTNYIGH-------KTKDLQAVCG 273

Qy         66 -KAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGEPTMEVRINVSRQQVEKVFGLEEYWC 124
                    :::| :  |                         :   :   : || |    |
Db        274 FSCDDLSKLFAEMKG---------------------LRTLVTTLKDQVTKETEKNELIA 311

Qy        125 QCVAWSSSGTTKSQKAYIRIARLRKNFEQ-------------EPLAKEVSLEQGIVLPC 170
              | |       |::    :    | || ::                 : ::||     ::||
Db        312 QIV-------TRTPGVCLHNGVLHKNRDEWTVDSCTECTCQNSATICRKVSCP---LMPC 361

Qy        171 RP---PEG--------IPPAEVEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANY 219
                  |:|            |: :|     |      |        :|   |  : ::  |   |: |
Db        362 TNATIPDGECCPRCWPSDSADDDWSPWSDWTPCS------VTCGHG-IQQRGRSCDSLNN 414

Qy        220 TCVAKNIVAR--------RRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCTN 271
               |   ::  |         :|            :|||  |: || || :|| |   | | | :
Db        415 PCEGSSVQTRSCQIQDCDKRFKQ-------DGGWSHWSPWSSCSVTCGSGQITRIRLCNS 467

Qy        272 PAPLNGGAFCEGQNVHDRTVSSLLVSVDGSWSPWSKWSACGLDC---THWRSRECSDPAP 328
              | |   |  |||:   ::           ::| | ||| |   | :  |        | | |::| |
Db        468 PVPQLNGKQCEGEGRENKPCQKDPCPINGQWGPWSLWDTCTVTCGGGMQKRERLCNNPKP 527

Qy        329 RNGGEECQGTDLDTRNCTSDLC 350
               :  |::| |   |:: |     |
Db        528 QYEGKDCIGEPTDSQICNKQDC 549


RESULT 9
SM5B_MOUSE
ID   SM5B_MOUSE     STANDARD;      PRT;  1093 AA.
AC   Q60519;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Semaphorin 5B precursor (Semaphorin G) (Sema G).
GN   SEMA5B OR SEMAG OR SEMG.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NMRI;
RX   MEDLINE=96414430; PubMed=8817451;
RA   Adams R.H., Betz H., Pueschel A.W.;
RT   "A novel class of murine semaphorins with homology to thrombospondin
RT   is differentially expressed during early embryogenesis.";
RL   Mech. Dev. 57:33-45(1996).
CC   -!- FUNCTION: May act as positive axonal guidance cues.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- TISSUE SPECIFICITY: In adult, only detected in brain.
CC   -!- DEVELOPMENTAL STAGE: Differentially expressed in embryonic and
CC       adult tissues. Its abundance decreases from E10 to birth.
CC   -!- SIMILARITY: Belongs to the semaphorin family.
CC   -!- SIMILARITY: Contains 1 Sema domain.
CC   -!- SIMILARITY: Contains 7 TSP type-1 domains.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X97818; CAA66398.1; -.
DR   MGD; MGI:107555; Sema5b.
DR   InterPro; IPR003659; Plexin-like.
DR   InterPro; IPR002165; Plexin_repeat.
DR   InterPro; IPR001627; Sema.
DR   InterPro; IPR000884; TSP1.
DR   InterPro; IPR008085; TSP_1.
DR   Pfam; PF01437; PSI; 1.
DR   Pfam; PF01403; Sema; 1.
DR   Pfam; PF00090; tsp_1; 5.
DR   PRINTS; PR01705; TSP1REPEAT.
DR   SMART; SM00423; PSI; 1.
DR   SMART; SM00630; Sema; 1.
DR   SMART; SM00209; TSP1; 4.
DR   PROSITE; PS50092; TSP1; 5.
KW   Signal; Transmembrane; Repeat; Multigene family; Neurogenesis;
KW   Developmental protein; Glycoprotein.
FT   SIGNAL        1     19       POTENTIAL.
FT   CHAIN        20   1093       SEMAPHORIN 5B.
FT   DOMAIN       20    978       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    979    999       POTENTIAL.
FT   DOMAIN     1000   1093       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      236    518       SEMA.
FT   DOMAIN      551    605       TSP TYPE-1 1.
FT   DOMAIN      606    662       TSP TYPE-1 2.
FT   DOMAIN      664    713       TSP TYPE-1 3.
FT   DOMAIN      721    776       TSP TYPE-1 4.
FT   DOMAIN      795    850       TSP TYPE-1 5.
FT   DOMAIN      852    907       TSP TYPE-1 6.
FT   DOMAIN      908    952       TSP TYPE-1 7.
FT   CARBOHYD     59     59       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     95     95       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    157    157       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    178    178       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    287    287       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    333    333       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    378    378       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    532    532       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    539    539       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    547    547       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    602    602       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    728    728       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    944    944       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   1093 AA;  120326 MW;  29E5C9B1E8108717 CRC64;

  Query Match              5.4%;  Score 258.5;  DB 1;  Length 1093;
  Best Local Similarity   22.2%;  Pred. No. 1e-11;
  Matches  104;  Conservative   31;  Mismatches  139;  Indels  195;  Gaps   16;

Qy        163 EQGIVLPCRPPEGIPPAEVEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCV 222
              ||      |||                   | ||                | |   :  |    |   ||
Db        737 EQRFRFTCRAP--------------LPDP-----------HGLQFGKRR---TETRTCP 767
```

```
Qy        223 AKNIVA-----------RRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCTN 271
              |     |          | |  |    :    ||||:||  || ||   |   |:: | |:|||
Db        768 ADGTGACDTDALVEDLLRSGSTSPHTL---NGGWATWGPWSSCSRDCELGFRVRKRTCTN 824

Qy        272 PAPLNGGAFCEGQNVHDRTVSSLLVSVDGSWSPWSKWSACGLDC---THWRSRECSDPAP 328
              | | ||| | |        :   :     | |:|| |: || |    |      : |:| |: |||
Db        825 PEPRNGGLPCVGDAAEYQDCNPQACPVRGAWSCWTAWSQCSASCGGGHYQRTRSCTSPAP 884

Qy        329 RNGGEECQGTDLDTRNCTSDLCVHSASGPEDVALYVGLIAVAVCLVLLLLVLILVYCRKK 388
                  | : | |   :   |::  |       || :|:      ||                  :
Db        885 SPGEDICLGLHTEEALCSTQAC------PEGWSLW---SEWGVC--------------TE 921

Qy        389 EGLDSDVADSSILTSGFQPVSIKPSKADNPHLLTIQPDLSTTTTYQGSLCPRQDGPSPKF 448
               :|  |       |  |                       |      :  |    || :  |
Db        922 DGAQSRSRSCEELLPG-------------------PGACVGNSSQSRPCPYSEIPV--- 958

Qy        449 QLTNGHLLSPLGGGRHTLHHSSPTSEAEEFVSRLSTQNYFRSLPRGTSNMTYGTFNFLGG 508
                                       | |  ||  |                     | ||
Db        959 -------------------ILPASSVEETTS------------------CGGFNL--- 976

Qy        509 RLMIPNTGISLLIPPDAIPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLL 568
                                                    : |  :  |       |||      ||
Db        977 -----------------------------------IHLIVTG--------VSCFLVSGLL 993

Qy        569 TRPVILAMDHCGEPSPDSWSLRLKKQSCEGSWEQDVLHLGEEAPSHLYY 617
              |  | |:  ||   | :|                  ::|      |:||:|
Db        994 TLAVYLSCQHCQRQSQES----------------TLVH--PATPNHLHY 1024


RESULT 10
BAI2_HUMAN
ID   BAI2_HUMAN     STANDARD;      PRT;  1572 AA.
AC   O60241;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Brain-specific angiogenesis inhibitor 2 precursor.
GN   BAI2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fetal brain;
RX   MEDLINE=98194217; PubMed=9533023;
RA   Shiratsuchi T., Nishimori H., Ichise H., Nakamura Y., Tokino T.;
RT   "Cloning and characterization of BAI2 and BAI3, novel genes homologous
RT   to brain-specific angiogenesis inhibitor 1 (BAI1).";
RL   Cytogenet. Cell Genet. 79:103-108(1997).
CC   -!- FUNCTION: MIGHT BE INVOLVED IN ANGIOGENESIS INHIBITION.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: STRONGLY EXPRESSED IN BRAIN. ALSO DETECTED IN
CC       HEART, THYMUS, SKELETAL MUSCLE, AND DIFFERENT CELL LINES.
CC   -!- SIMILARITY: Belongs to family 2 of G-protein coupled receptors.
CC   -!- SIMILARITY: Contains 4 TSP type-1 domains.
CC   -!- SIMILARITY: Contains 1 GPS domain.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AB005298; BAA25362.1; -.
DR   PIR; T00027; T00027.
DR   Genew; HGNC:944; BAI2.
DR   MIM; 602683; -.
DR   InterPro; IPR000832; GPCR_secretin.
DR   InterPro; IPR001879; hormn_receptor.
DR   InterPro; IPR000203; PKD_cys_rich.
DR   InterPro; IPR000884; TSP1.
DR   Pfam; PF00002; 7tm_2; 1.
DR   Pfam; PF01825; GPS; 1.
DR   Pfam; PF02793; HRM; 1.
DR   Pfam; PF00090; tsp_1; 4.
DR   SMART; SM00303; GPS; 1.
DR   SMART; SM00008; HormR; 1.
DR   SMART; SM00209; TSP1; 4.
DR   PROSITE; PS50221; GPS; 1.
DR   PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; FALSE_NEG.
DR   PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; FALSE_NEG.
DR   PROSITE; PS50227; G_PROTEIN_RECEP_F2_3; 1.
DR   PROSITE; PS50261; G_PROTEIN_RECEP_F2_4; 1.
DR   PROSITE; PS50092; TSP1; 4.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW   Repeat.
FT   SIGNAL        1     20       POTENTIAL.
FT   CHAIN        21   1572       BRAIN-SPECIFIC ANGIOGENESIS INHIBITOR 2.
FT   DOMAIN       21    924       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    925    945       1 (POTENTIAL).
FT   DOMAIN      946    953       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    954    974       2 (POTENTIAL).
FT   DOMAIN      975    982       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    983   1003       3 (POTENTIAL).
FT   DOMAIN     1004   1024       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   1025   1045       4 (POTENTIAL).
FT   DOMAIN     1046   1066       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   1067   1087       5 (POTENTIAL).
FT   DOMAIN     1088   1141       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   1142   1162       6 (POTENTIAL).
FT   DOMAIN     1163   1168       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   1169   1189       7 (POTENTIAL).
FT   DOMAIN     1190   1572       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      297    350       TSP TYPE-1 1.
FT   DOMAIN      352    405       TSP TYPE-1 2.
FT   DOMAIN      407    460       TSP TYPE-1 3.
FT   DOMAIN      463    516       TSP TYPE-1 4.
FT   DOMAIN      859    911       GPS.
FT   DOMAIN      117    122       POLY-GLU.
FT   DOMAIN      177    180       POLY-ASN.
FT   DOMAIN      222    225       POLY-THR.
FT   DOMAIN     1303   1306       POLY-PRO.
FT   DOMAIN     1352   1358       POLY-GLY.
FT   DOMAIN     1413   1418       POLY-PRO.
FT   CARBOHYD     94     94       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    179    179       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    180    180       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    344    344       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    425    425       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    548    548       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    633    633       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    855    855       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   1572 AA;  171140 MW;  A9775645B77BC285 CRC64;

  Query Match             5.4%;  Score 258;  DB 1;  Length 1572;
  Best Local Similarity   18.8%;  Pred. No. 1.9e-11;
  Matches  172;  Conservative  110;  Mismatches  312;  Indels  322;  Gaps   37;

Qy        173 PEGIPPAEVEWLRNEDLVDPSLDPNVY-------------------ITREHSLVVRQARL 213
              ||  |  : :| |: |            :| :|                :|     | ||  |
Db        271 PEEEPKVKTQWPRSAD------EPGLYMAQTGDPAAEEWSPWSVCSLTCGQGLQVR-TRS 323

Qy        214 ADTANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPA 273
                :: |  :    :    |   : :       |:|  |    |  ||:||  |||||  :  |  |:|
Db        324 CVSSPYGTLCSGPLRETRPCNNSATCPVHGVWEEWGSWSLCSRSCGRGSRSRMRTCV--P 381

Qy        274 PLNGGAFCEGQNVHDRTVSSLLVSVDGSWSPWSKWSACGLDC---THWRSRECSDPAPR- 329
              | :||  |||  :  :   |      |:|  |   |   |    |   |    |||:||    |
Db        382 PQHGGKACEGPELQTKLCSMAACPVEGQWLEWGPWGPCSTSCANGTQQRSRKCSVAGPAW 441

Qy        330 -----------------------------------------------------NGGEECQ 336
                                                                         |  |:
```

```
Db     442 ATCTGALTDTRECSNLECPATDSKWGPWNAWSLCSKTCDTGWQRRFRMCQATGTQGYPCE 501

Qy     337 GTDLDTRNCTSDLC--VHSASGPEDVAL-------------------------------- 362
           ||  : : |:    |    |      | | |
Db     502 GTGEEVKPCSEKRCPAFHEMCRDEYVMLMTWKKAAAGEIIYNKCPPNASGSASRRCLLSA 561

Qy     363 ----YVGLIAVAVCL---VLLLLVLILVYCRKKEGLDSDVADSSILTSGFQPVSIKPSKA 415
               | || : | |:        | : :  :   | : : :    | :: |   | :  : :
Db     562 QGVAYWGLPSFARCISHEYRYLYLSLREHLAKGQRMLAGEGMSQVVRS-LQELLARRTYY 620

Qy     416 DNPHLLTIQPDLSTTTTY-QGSLCPRQDGPSPKFQLT------------NGHLLSPLGG 461
               | ::    : | |: : :   |   |      ||:              :    :||  |
Db     621 SGDLLFSVDILRNVTDTFKRATYVPSADDVQRFFQVVSFMVDAENKEKWDDAQQVSP--G 678

Qy     462 GRHTLHHSSPTSEAEEFV------------SRLSTQNYFRSLPRG-----TSNMTYGTFN 504
              | |          |:|:            | :  | |   |: |       :|::|:
Db     679 SVHLLR------VVEDFIHLVGDALKAFQSSLIVTDNLVISIQREPVSAVSSDITFPMRG 732

Qy     505 FLG---------GRLMIPNTGISLLIP-----------------PDAIPRGK-------- 530
             |             || :|    :||  |                    |  :| |
Db     733 RRGMKDWVRHSEDRLFLPKEVLSLSSPGKPATSGAAGSPGRGRGPGTVPPGPGHSHQRLL 792

Qy     531 ---------------IYE-IYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVIL 574
                          :|  : | |  |    |  |||    :::  |:  ||      |:|
Db     793 PADPDESSYFVIGAVLYRTLGLILPPP---RPPLAVTSRVMT--VTVRPPTQPPAEPLIT 847

Qy     575 A-----MDHCGEPSPDSWSLRLKKQSCEGSWEQDVLHLGEEAPSHLYYCQLEASACYVFT 629
                 ::   :|   ||     :   :   | |: :        |    :|    ||
Db     848 VELSYIINGTTDPHCASWDYS-RADASSGDWDTENCQTLETQAAHT-RCQ---------C 896

Qy     630 EQLGRFALVGE-----ALSVAAAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQ 684
           : |   ||:: :       |:|  :   : |::    |:| :|    : :|          |       ::
Db     897 QHLSTFAVLAQPPKDLTLELAGSPSVPLVIGCAVSCMALLTLLAIYA------AFWRFIK 950

Qy     685 LEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWKSKLLVSYQEIPFYHIWNGTQRY 744
            |: :                     |  |||  :  |::     :||      :       |    :
Db     951 SERSI---------------ILLNFCLSI--LASNI---LILVGQSRVLSKGVCTMTAAF 990

Qy     745 LHCTFTLERVSPSTSDLACKLWV---------------------------WQVEGDGQS 776
            ||  |              |:   ||                           | :   :
Db     991 LHFFF-----------LSSFCWVLTEAWQSYLAVIGRMRTRLVRKRFLCLGWGLPALVVA 1039

Qy     777 FSINFNITKDTRFAELLALESEAG-VPALVGPSA------FKIPFLIRQKI-----ISSL 824
            |: |  ||      :    |  |  |  : |  |||:|          |  ::  |:      ||
Db    1040 VSVGFTRTKGYGTSSYCWLSLEGGLLYAFVGPAAVIVLVNMLIGIIVFNKLMARDGISDK 1099

Qy     825 DPPCRRGAD---WRTL 837
              | |::     | :|
Db    1100 SKKQRAGSERCPWASL 1115
```

```
RESULT 11
SM5B_HUMAN
ID   SM5B_HUMAN     STANDARD;      PRT;  1093 AA.
AC   Q9P283;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Semaphorin 5B precursor.
GN   SEMA5B OR KIAA1445.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=20277482; PubMed=10819331;
RA   Nagase T., Kikuno R., Ishikawa K.-I., Hirosawa M., Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes. XVII.
RT   The complete sequences of 100 new cDNA clones from brain which code
RT   for large proteins in vitro.";
RL   DNA Res. 7:143-150(2000).
CC   -!- FUNCTION: May act as positive axonal guidance cues (By
CC       similarity).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- SIMILARITY: Belongs to the semaphorin family.
CC   -!- SIMILARITY: Contains 1 Sema domain.
CC   -!- SIMILARITY: Contains 7 TSP type-1 domains.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AB040878; BAA95969.1; ALT_INIT.
DR   Genew; HGNC:10737; SEMA5B.
DR   InterPro; IPR003659; Plexin-like.
DR   InterPro; IPR002165; Plexin_repeat.
DR   InterPro; IPR001627; Sema.
DR   InterPro; IPR000884; TSP1.
DR   InterPro; IPR008085; TSP_1.
DR   Pfam; PF01437; PSI; 1.
DR   Pfam; PF01403; Sema; 1.
DR   Pfam; PF00090; tsp_1; 5.
DR   PRINTS; PR01705; TSP1REPEAT.
DR   SMART; SM00423; PSI; 1.
DR   SMART; SM00630; Sema; 1.
DR   SMART; SM00209; TSP1; 4.
DR   PROSITE; PS50092; TSP1; 5.
KW   Signal; Transmembrane; Repeat; Multigene family; Neurogenesis;
KW   Developmental protein; Glycoprotein.
FT   SIGNAL        1     26       POTENTIAL.
FT   CHAIN        27   1093       SEMAPHORIN 5B.
FT   DOMAIN       20    978       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    979    999       POTENTIAL.
FT   DOMAIN     1000   1093       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      236    518       SEMA.
FT   DOMAIN      551    605       TSP TYPE-1 1.
FT   DOMAIN      606    662       TSP TYPE-1 2.
FT   DOMAIN      664    713       TSP TYPE-1 3.
FT   DOMAIN      721    776       TSP TYPE-1 4.
FT   DOMAIN      795    850       TSP TYPE-1 5.
FT   DOMAIN      852    907       TSP TYPE-1 6.
FT   DOMAIN      908    952       TSP TYPE-1 7.
FT   CARBOHYD     59     59      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     95     95      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    157    157      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    178    178      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    287    287      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    333    333      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    378    378      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    532    532      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    539    539      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    547    547      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    602    602      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    728    728      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    944    944      N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   1093 AA;  119866 MW;  F1FDEFB87CEAF0EF CRC64;

  Query Match             5.3%;  Score 252;  DB 1;  Length 1093;
  Best Local Similarity   29.4%;  Pred. No. 3.3e-11;
  Matches   68;  Conservative   34;  Mismatches   85;  Indels   44;  Gaps   10;

Qy     241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDG 300
           | | || || || |||||| |   :|:||||:|||   |     | |  :|          :
Db     851 VRGAWSCWTSWSPCSASCGGGHYQRTRSCTSPAPSPGEDICLG--LHTEEALCATQACPE 908

Qy     301 SWSPWSKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCT-SDL-CVHSASGPE 358
           |||||:|| |  |       ||| | :  |  |   ||     :| |  |:: :  || |
Db     909 GWSPWSEWSKCTDDGAQSRSRHCEELLP--GSSACAGNSSQSRPCPYSEIPVILPASSME 966

Qy     359 DVALYVG-----LIAVAVCLVL---LLLVLILVYCR--KKEGLDSDVADSSILTSGFQPV 408
```

```
                    :      |       |:|   :     |      ||  :  :  :  |:   :::   :|  :
Db        967 EATGCAGFNLIHLVATGISCFLGSGLLTLAVYLSCQHCQRQSQESTL------------- 1013

Qy        409 SIKPSKADNPHLLTIQPDLSTTTTYQGSLCPRQDGPSP-KFQLTNGHLLSP 458
                 :  |:   ::  |                |:|     |:  :   :|  :|:    |  :  |  |
Db       1014 -VHPATPNHLH-------------YKGGGTPKNEKYTPMEFKTLNKNNLIP 1050


RESULT 12
TSP1_MOUSE
ID   TSP1_MOUSE     STANDARD;      PRT;  1170 AA.
AC   P35441;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Thrombospondin 1 precursor.
GN   THBS1 OR TSP1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92128941; PubMed=1774063;
RA   Lawler J., Duquette M., Ferro P., Copeland N.G., Gilbert D.J.,
RA   Jenkins N.A.;
RT   "Characterization of the murine thrombospondin gene.";
RL   Genomics 11:587-600(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92147683; PubMed=1371115;
RA   Laherty C.D., O'Rourke K., Wolf F.W., Katz R., Seldin M.F.,
RA   Dixit V.M.;
RT   "Characterization of mouse thrombospondin 2 sequence and expression
RT   during cell growth and development.";
RL   J. Biol. Chem. 267:3274-3281(1992).
RN   [3]
RP   SEQUENCE OF 1-490 FROM N.A.
RX   MEDLINE=90375546; PubMed=2398070;
RA   Bornstein P., Alfi D., Devarayalu S., Framson P., Li P.;
RT   "Characterization of the mouse thrombospondin gene and evaluation of
RT   the role of the first intron in human gene expression.";
RL   J. Biol. Chem. 265:16691-16698(1990).
CC   -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC       cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC       laminin, type V collagen and integrins alpha-V/beta-1, alpha-
CC       V/beta-3 and alpha-IIb/beta-3.
CC   -!- SUBUNIT: Homotrimer; disulfide-linked.
CC   -!- SIMILARITY: Belongs to the thrombospondin family.
CC   -!- SIMILARITY: Contains 1 VWFC domain.
CC   -!- SIMILARITY: Contains 3 EGF-like domains.
CC   -!- SIMILARITY: Contains 3 TSP type-1 domains.
CC   -!- SIMILARITY: Contains 7 TSP type-3 domains.
CC   -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M62470; AAA50611.1; -.
DR   EMBL; M62450; AAA50611.1; JOINED.
DR   EMBL; M62451; AAA50611.1; JOINED.
DR   EMBL; M62452; AAA50611.1; JOINED.
DR   EMBL; M62453; AAA50611.1; JOINED.
DR   EMBL; M62454; AAA50611.1; JOINED.
DR   EMBL; M62455; AAA50611.1; JOINED.
DR   EMBL; M62456; AAA50611.1; JOINED.
DR   EMBL; M62457; AAA50611.1; JOINED.
DR   EMBL; M62458; AAA50611.1; JOINED.
DR   EMBL; M62459; AAA50611.1; JOINED.
DR   EMBL; M62460; AAA50611.1; JOINED.
DR   EMBL; M62461; AAA50611.1; JOINED.
DR   EMBL; M62462; AAA50611.1; JOINED.
DR   EMBL; M62463; AAA50611.1; JOINED.
DR   EMBL; M62464; AAA50611.1; JOINED.
DR   EMBL; M62465; AAA50611.1; JOINED.
DR   EMBL; M62466; AAA50611.1; JOINED.
DR   EMBL; M62467; AAA50611.1; JOINED.
DR   EMBL; M62468; AAA50611.1; JOINED.
DR   EMBL; M62469; AAA50611.1; JOINED.
DR   EMBL; M87276; AAA53063.1; -.
DR   EMBL; J05606; AAA40431.1; -.
DR   EMBL; J05605; AAA40431.1; JOINED.
DR   PIR; A40558; A40558.
DR   MGD; MGI:98737; Thbs1.
DR   GO; GO:0005615; C:extracellular space; IDA.
DR   GO; GO:0016525; P:negative regulation of angiogenesis; IDA.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR000884; TSP1.
DR   InterPro; IPR008085; TSP_1.
DR   InterPro; IPR003367; tsp_3.
DR   InterPro; IPR008859; TSPC.
DR   InterPro; IPR003129; TSPN.
DR   InterPro; IPR001007; VWF_C.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00090; tsp_1; 3.
DR   Pfam; PF02412; tsp_3; 13.
DR   Pfam; PF05735; TSPC; 1.
DR   Pfam; PF02210; TSPN; 1.
DR   Pfam; PF00093; vwc; 1.
DR   PRINTS; PR01705; TSP1REPEAT.
DR   SMART; SM00181; EGF; 3.
DR   SMART; SM00209; TSP1; 3.
DR   SMART; SM00210; TSPN; 1.
DR   SMART; SM00214; VWC; 1.
DR   PROSITE; PS00022; EGF_1; FALSE_NEG.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS50026; EGF_3; 2.
DR   PROSITE; PS50092; TSP1; 3.
DR   PROSITE; PS01208; VWFC_1; 1.
DR   PROSITE; PS50184; VWFC_2; 1.
KW   Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat;
KW   EGF-like domain; Signal.
FT   SIGNAL        1     18       POTENTIAL.
FT   CHAIN        19   1170       THROMBOSPONDIN 1.
FT   DOMAIN       19    232       HEPARIN-BINDING (POTENTIAL).
FT   DOMAIN       24    221       TSP N-TERMINAL.
FT   DOMAIN      316    373       VWFC.
FT   DOMAIN      379    429       TSP TYPE-1 1.
FT   DOMAIN      435    490       TSP TYPE-1 2.
FT   DOMAIN      492    547       TSP TYPE-1 3.
FT   DOMAIN      549    587       EGF-LIKE 1.
FT   DOMAIN      588    645       EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      646    690       EGF-LIKE 3.
FT   DOMAIN      723    758       TSP TYPE-3 1.
FT   DOMAIN      759    781       TSP TYPE-3 2.
FT   DOMAIN      782    817       TSP TYPE-3 3.
FT   DOMAIN      818    840       TSP TYPE-3 4.
FT   DOMAIN      841    878       TSP TYPE-3 5.
FT   DOMAIN      879    914       TSP TYPE-3 6.
FT   DOMAIN      915    950       TSP TYPE-3 7.
FT   DOMAIN      951   1170       C-TERMINAL.
FT   SITE        926    928       CELL ATTACHMENT SITE (POTENTIAL).
FT   DISULFID    270    270       INTERCHAIN (PROBABLE).
FT   DISULFID    274    274       INTERCHAIN (PROBABLE).
FT   DISULFID    391    423       BY SIMILARITY.
FT   DISULFID    395    428       BY SIMILARITY.
FT   DISULFID    406    413       BY SIMILARITY.
FT   DISULFID    447    484       BY SIMILARITY.
FT   DISULFID    451    489       BY SIMILARITY.
```

```
FT   DISULFID    462    474       BY SIMILARITY.
FT   DISULFID    504    541       BY SIMILARITY.
FT   DISULFID    508    546       BY SIMILARITY.
FT   DISULFID    519    531       BY SIMILARITY.
FT   DISULFID    551    562       BY SIMILARITY.
FT   DISULFID    556    572       BY SIMILARITY.
FT   DISULFID    575    586       BY SIMILARITY.
FT   DISULFID    592    608       BY SIMILARITY.
FT   DISULFID    599    617       BY SIMILARITY.
FT   DISULFID    620    644       BY SIMILARITY.
FT   DISULFID    650    663       BY SIMILARITY.
FT   DISULFID    657    676       BY SIMILARITY.
FT   DISULFID    678    689       BY SIMILARITY.
FT   DISULFID    705    713       BY SIMILARITY.
FT   DISULFID    718    738       BY SIMILARITY.
FT   DISULFID    754    774       BY SIMILARITY.
FT   DISULFID    777    797       BY SIMILARITY.
FT   DISULFID    813    833       BY SIMILARITY.
FT   DISULFID    836    856       BY SIMILARITY.
FT   DISULFID    874    894       BY SIMILARITY.
FT   DISULFID    910    930       BY SIMILARITY.
FT   DISULFID    946   1167       BY SIMILARITY.
FT   CARBOHYD    248    248       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    360    360       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    708    708       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1067   1067       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT   1025   1025       F -> L (IN REF. 2).
SQ   SEQUENCE   1170 AA;  129646 MW;  0443E493615E7F06 CRC64;

  Query Match             5.2%;  Score 248;  DB 1;  Length 1170;
  Best Local Similarity   29.9%;  Pred. No. 7.4e-11;
  Matches   53;  Conservative   24;  Mismatches   76;  Indels   24;  Gaps    4;

Qy        207 VVRQARLADTANYTCVAKNIVAR--------RRSASAAVIVYVNGGWSTWTEWSVCSASC 258
              : ::  |  |: |   |    ::  |              :|          :||||  |:  || ||  :|
Db        399 IQQRGRSCDSLNNRCEGSSVQTRTCHIQECDKRFKQ-------DGGWSHWSPWSSCSVTC 451

Qy        259 GRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDGSWSPWSKWSACGLDC--- 315
               | |    | | | :|:|    |  |||:      :          ::| |  ||| |    | : |
Db        452 GDGVITRIRLCNSPSPQMNGKPCEGEARETKACKKDACPINGGWGPWSPWDICSVTCGGG 511

Qy        316 THWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCVHSASGPEDVALYVGLIAVAVC 372
                 ||| |::| |: ||::| |     : : |      |           | |   |        | | |
Db        512 VQRRSRLCNNPTPQFGGKDCVGDVTENQVCNKQDC------PIDGCLSNPCFAGAKC 562

RESULT 13
TSP1_HUMAN
ID   TSP1_HUMAN     STANDARD;      PRT;  1170 AA.
AC   P07996; Q15667;
DT   01-AUG-1988 (Rel. 08, Created)
DT   01-AUG-1988 (Rel. 08, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Thrombospondin 1 precursor.
GN   THBS1 OR TSP1 OR TSP.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Endothelial cells;
RX   MEDLINE=87057617; PubMed=2430973;
RA   Lawler J., Hynes R.O.;
RT   "The structure of human thrombospondin, an adhesive glycoprotein with
RT   multiple calcium-binding sites and homologies with several different
RT   proteins.";
RL   J. Cell Biol. 103:1635-1648(1986).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89139590; PubMed=2918029;
RA   Hennessy S.W., Frazier B.A., Kim D.D., Deckwerth T.L.,
RA   Baumgartel D.M., Rotwein P., Frazier W.A.;
RT   "Complete thrombospondin mRNA sequence includes potential regulatory
RT   sites in the 3' untranslated region.";
RL   J. Cell Biol. 108:729-736(1989).
RN   [3]
RP   SEQUENCE OF 1-397 FROM N.A.
RX   MEDLINE=87157592; PubMed=3030396;
RA   Kobayashi S., Eden-Mccutchan F., Framson P., Bornstein P.;
RT   "Partial amino acid sequence of human thrombospondin as determined by
RT   analysis of cDNA clones: homology to malarial circumsporozoite
RT   proteins.";
RL   Biochemistry 25:8418-8425(1986).
RN   [4]
RP   SEQUENCE OF 1-374 FROM N.A.
RX   MEDLINE=86287276; PubMed=3461443;
RA   Dixit V.M., Hennessy S.W., Grant G.A., Rotwein P., Frazier W.A.;
RT   "Characterization of a cDNA encoding the heparin and collagen binding
RT   domains of human thrombospondin.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:5449-5453(1986).
RN   [5]
RP   SEQUENCE OF 1-166 FROM N.A.
RX   MEDLINE=89291870; PubMed=2544587;
RA   Laherty C.D., Gierman T.M., Dixit V.M.;
RT   "Characterization of the promoter region of the human thrombospondin
RT   gene. DNA sequences within the first intron increase transcription.";
RL   J. Biol. Chem. 264:11222-11227(1989).
RN   [6]
RP   SEQUENCE OF 1028-1170 FROM N.A.
RA   la Fleur M., Jobin C., Gauthier J., Kreis C.G.;
RL   Submitted (XXX-1992) to the EMBL/GenBank/DDBJ databases.
RN   [7]
RP   CARBOHYDRATE-LINKAGE SITES TRP-385; SER-394; TRP-438; TRP-441;
RP   THR-450; TRP-498 AND THR-507.
RC   TISSUE=Platelet;
RX   MEDLINE=21125860; PubMed=11067851;
RA   Hofsteenge J., Huwiler K.G., Macek B., Hess D., Lawler J.,
RA   Mosher D.F., Peter-Katalinic J.;
RT   "C-mannosylation and O-fucosylation of the thrombospondin type 1
RT   module.";
RL   J. Biol. Chem. 276:6485-6498(2001).
RN   [8]
RP   THROMBOSPONDIN DOMAIN DISULFIDE BRIDGES.
RX   MEDLINE=22338361; PubMed=12450399;
RA   Huwiler K.G., Vestling M.M., Annis D.S., Mosher D.F.;
RT   "Biophysical characterization, including disulfide bond assignments,
RT   of the anti-angiogenic type 1 domains of human thrombospondin-1.";
RL   Biochemistry 41:14329-14339(2002).
CC   -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC       cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC       laminin, type V collagen and integrins alpha-V/beta-1, alpha-
CC       V/beta-3 and alpha-IIb/beta-3.
CC   -!- SUBUNIT: Homotrimer; disulfide-linked.
CC   -!- SIMILARITY: Belongs to the thrombospondin family.
CC   -!- SIMILARITY: Contains 1 VWFC domain.
CC   -!- SIMILARITY: Contains 3 EGF-like domains.
CC   -!- SIMILARITY: Contains 3 TSP type-1 domains.
CC   -!- SIMILARITY: Contains 7 TSP type-3 domains.
CC   -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M25631; AAA36741.1; -.
DR   EMBL; X04665; CAA28370.1; -.
DR   EMBL; X14787; CAA32889.1; -.
DR   EMBL; M14326; AAA61237.1; ALT_SEQ.
DR   EMBL; J04835; AAA61178.1; -.
DR   EMBL; M99425; AAB59366.1; -.
DR   PIR; A26155; TSHUP1.
```

```
DR   PDB; 1LSL; 18-DEC-02.
DR   GlycoSuiteDB; P07996; -.
DR   Genew; HGNC:11785; THBS1.
DR   MIM; 188060; -.
DR   GO; GO:0004866; F:endopeptidase inhibitor activity; TAS.
DR   GO; GO:0004871; F:signal transducer activity; TAS.
DR   GO; GO:0007275; P:development; TAS.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR000884; TSP1.
DR   InterPro; IPR008085; TSP_1.
DR   InterPro; IPR003367; tsp_3.
DR   InterPro; IPR008859; TSPC.
DR   InterPro; IPR003129; TSPN.
DR   InterPro; IPR001007; VWF_C.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00090; tsp_1; 3.
DR   Pfam; PF02412; tsp_3; 13.
DR   Pfam; PF05735; TSPC; 1.
DR   Pfam; PF02210; TSPN; 1.
DR   Pfam; PF00093; vwc; 1.
DR   PRINTS; PR01705; TSP1REPEAT.
DR   SMART; SM00181; EGF; 3.
DR   SMART; SM00209; TSP1; 3.
DR   SMART; SM00210; TSPN; 1.
DR   SMART; SM00214; VWC; 1.
DR   PROSITE; PS00022; EGF_1; FALSE_NEG.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS50026; EGF_3; 2.
DR   PROSITE; PS50092; TSP1; 3.
DR   PROSITE; PS01208; VWFC_1; 1.
DR   PROSITE; PS50184; VWFC_2; 1.
KW   Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat;
KW   EGF-like domain; Signal; 3D-structure.
FT   SIGNAL        1     18
FT   CHAIN        19   1170       THROMBOSPONDIN 1.
FT   DOMAIN       19    232       HEPARIN-BINDING (POTENTIAL).
FT   DOMAIN       24    221       TSP N-TERMINAL.
FT   DOMAIN      316    373       VWFC.
FT   DOMAIN      379    429       TSP TYPE-1 1.
FT   DOMAIN      435    490       TSP TYPE-1 2.
FT   DOMAIN      492    547       TSP TYPE-1 3.
FT   DOMAIN      549    587       EGF-LIKE 1.
FT   DOMAIN      588    645       EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      646    690       EGF-LIKE 3.
FT   DOMAIN      723    758       TSP TYPE-3 1.
FT   DOMAIN      759    781       TSP TYPE-3 2.
FT   DOMAIN      782    817       TSP TYPE-3 3.
FT   DOMAIN      818    840       TSP TYPE-3 4.
FT   DOMAIN      841    878       TSP TYPE-3 5.
FT   DOMAIN      879    914       TSP TYPE-3 6.
FT   DOMAIN      915    950       TSP TYPE-3 7.
FT   DOMAIN      951   1170       C-TERMINAL.
FT   SITE        926    928       CELL ATTACHMENT SITE (POTENTIAL).
FT   DISULFID    270    270       INTERCHAIN (PROBABLE).
FT   DISULFID    274    274       INTERCHAIN (PROBABLE).
FT   DISULFID    391    423
FT   DISULFID    395    428
FT   DISULFID    406    413
FT   DISULFID    447    484
FT   DISULFID    451    489
FT   DISULFID    462    474
FT   DISULFID    504    541
FT   DISULFID    508    546
FT   DISULFID    519    531
FT   DISULFID    551    562       BY SIMILARITY.
FT   DISULFID    556    572       BY SIMILARITY.
FT   DISULFID    575    586       BY SIMILARITY.
FT   DISULFID    592    608       BY SIMILARITY.
FT   DISULFID    599    617       BY SIMILARITY.
FT   DISULFID    620    644       BY SIMILARITY.
FT   DISULFID    650    663       BY SIMILARITY.
FT   DISULFID    657    676       BY SIMILARITY.
FT   DISULFID    678    689       BY SIMILARITY.
FT   DISULFID    705    713       BY SIMILARITY.
FT   DISULFID    718    738       BY SIMILARITY.
FT   DISULFID    754    774       BY SIMILARITY.
FT   DISULFID    777    797       BY SIMILARITY.
FT   DISULFID    813    833       BY SIMILARITY.
FT   DISULFID    836    856       BY SIMILARITY.
FT   DISULFID    874    894       BY SIMILARITY.
FT   DISULFID    910    930       BY SIMILARITY.
FT   DISULFID    946   1167       BY SIMILARITY.
FT   CARBOHYD    248    248       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    360    360       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    385    385       C-LINKED (MAN).
FT                                /FTId=CAR_000205.
FT   CARBOHYD    394    394       O-LINKED (FUC. . .).
FT                                /FTId=CAR_000206.
FT   CARBOHYD    438    438       C-LINKED (MAN).
FT                                /FTId=CAR_000207.
FT   CARBOHYD    441    441       C-LINKED (MAN).
FT                                /FTId=CAR_000208.
FT   CARBOHYD    450    450       O-LINKED (FUC. . .).
FT                                /FTId=CAR_000209.
FT   CARBOHYD    498    498       C-LINKED (MAN).
FT                                /FTId=CAR_000210.
FT   CARBOHYD    507    507       O-LINKED (FUC. . .).
FT                                /FTId=CAR_000211.
FT   CARBOHYD    708    708       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1067   1067       N-LINKED (GLCNAC. . .) (POTENTIAL).

  Query Match             5.1%;  Score 246;  DB 1;  Length 1170;
  Best Local Similarity   30.3%;  Pred. No. 1.1e-10;
  Matches   47;  Conservative   24;  Mismatches   66;  Indels   18;  Gaps    3;

Qy        207 VVRQARLADTANYTCVAKNIVAR--------RRSASAAVIVYVNGGWSTWTEWSVCSASC 258
              : ::  |  |:  |  |    ::   |         :|          :||||  |:  || || :|
Db        399 IQQRGRSCDSLNNRCEGSSVQTRTCHIQECDKRFKQ-------DGGWSHWSPWSSCSVTC 451

Qy        259 GRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDGSWSPWSKWSACGLDC--- 315
              |  |    |  |  |  :|:|    |   |||:      :          ::|  |  |||  |    |  :  |
Db        452 GDGVITRIRLCNSPSPQMNGKPCEGEARETKACKKDACPINGGWGPWSPWDICSVTCGGG 511

Qy        316 THWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLC 350
                 |||  |::|  |:  ||::|  |      :  :  |      |
Db        512 VQKRSRLCNNPTPQFGGKDCVGDVTENQICNKQDC 546


RESULT 14
TSP2_CHICK
ID   TSP2_CHICK     STANDARD;      PRT;  1178 AA.
AC   P35440;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Thrombospondin 2 precursor.
GN   THBS2 OR TSP2.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91217026; PubMed=2022631;
RA   Lawler J., Duquette M., Ferro P.;
RT   "Cloning and sequencing of chicken thrombospondin.";
RL   J. Biol. Chem. 266:8039-8043(1991).
CC   -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC       cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC       laminin and type V collagen.
CC   -!- SUBUNIT: Homotrimer; disulfide-linked.
CC   -!- SIMILARITY: Belongs to the thrombospondin family.
CC   -!- SIMILARITY: Contains 1 VWFC domain.
```

```
CC   -!- SIMILARITY: Contains 3 EGF-like domains.
CC   -!- SIMILARITY: Contains 3 TSP type-1 domains.
CC   -!- SIMILARITY: Contains 7 TSP type-3 domains.
CC   -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M60853; AAA51437.1; -.
DR   PIR; A39804; A39804.
DR   HSSP; P00740; 1EDM.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR000884; TSP1.
DR   InterPro; IPR008085; TSP_1.
DR   InterPro; IPR003367; tsp_3.
DR   InterPro; IPR008859; TSPC.
DR   InterPro; IPR003129; TSPN.
DR   InterPro; IPR001007; VWF_C.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF00090; tsp_1; 3.
DR   Pfam; PF02412; tsp_3; 13.
DR   Pfam; PF05735; TSPC; 1.
DR   Pfam; PF02210; TSPN; 1.
DR   Pfam; PF00093; vwc; 1.
DR   PRINTS; PR01705; TSP1REPEAT.
DR   SMART; SM00181; EGF; 2.
DR   SMART; SM00209; TSP1; 3.
DR   SMART; SM00210; TSPN; 1.
DR   SMART; SM00214; VWC; 1.
DR   PROSITE; PS00022; EGF_1; FALSE_NEG.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS50026; EGF_3; 2.
DR   PROSITE; PS50092; TSP1; 3.
DR   PROSITE; PS01208; VWFC_1; 1.
DR   PROSITE; PS50184; VWFC_2; 1.
KW   Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat;
KW   EGF-like domain; Signal.
FT   SIGNAL        1     22       POTENTIAL.
FT   CHAIN        23   1178       THROMBOSPONDIN 2.
FT   DOMAIN       25    221       TSP N-TERMINAL.
FT   DOMAIN        ?    232       HEPARIN-BINDING (POTENTIAL).
FT   DOMAIN      324    381       VWFC.
FT   DOMAIN      387    437       TSP TYPE-1 1.
FT   DOMAIN      443    498       TSP TYPE-1 2.
FT   DOMAIN      500    555       TSP TYPE-1 3.
FT   DOMAIN      555    595       EGF-LIKE 1.
FT   DOMAIN      596    653       EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      654    698       EGF-LIKE 3.
FT   DOMAIN      731    766       TSP TYPE-3 1.
FT   DOMAIN      767    789       TSP TYPE-3 2.
FT   DOMAIN      790    825       TSP TYPE-3 3.
FT   DOMAIN      826    848       TSP TYPE-3 4.
FT   DOMAIN      849    886       TSP TYPE-3 5.
FT   DOMAIN      887    922       TSP TYPE-3 6.
FT   DOMAIN      923    958       TSP TYPE-3 7.
FT   DOMAIN      959   1178       C-TERMINAL.
FT   SITE        934    935       CELL ATTACHMENT SITE (POTENTIAL).
FT   DISULFID    399    431       BY SIMILARITY.
FT   DISULFID    403    436       BY SIMILARITY.
FT   DISULFID    414    421       BY SIMILARITY.
FT   DISULFID    455    492       BY SIMILARITY.
FT   DISULFID    459    497       BY SIMILARITY.
FT   DISULFID    470    482       BY SIMILARITY.
FT   DISULFID    512    549       BY SIMILARITY.
FT   DISULFID    516    554       BY SIMILARITY.
FT   DISULFID    527    539       BY SIMILARITY.
FT   DISULFID    559    570       BY SIMILARITY.
FT   DISULFID    564    580       BY SIMILARITY.
FT   DISULFID    583    594       BY SIMILARITY.
FT   DISULFID    600    616       BY SIMILARITY.
FT   DISULFID    607    625       BY SIMILARITY.
FT   DISULFID    628    652       BY SIMILARITY.
FT   DISULFID    658    671       BY SIMILARITY.
FT   DISULFID    665    684       BY SIMILARITY.
FT   DISULFID    686    697       BY SIMILARITY.
FT   DISULFID    713    721       BY SIMILARITY.
FT   DISULFID    726    746       BY SIMILARITY.
FT   DISULFID    762    782       BY SIMILARITY.
FT   DISULFID    785    805       BY SIMILARITY.
FT   DISULFID    821    841       BY SIMILARITY.
FT   DISULFID    844    864       BY SIMILARITY.
FT   DISULFID    882    902       BY SIMILARITY.
FT   DISULFID    918    938       BY SIMILARITY.
FT   DISULFID    954   1175       BY SIMILARITY.
FT   CARBOHYD    157    157       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    244    244       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    317    317       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    322    322       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    463    463       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    590    590       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    716    716       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1075   1075       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   1178 AA;  131816 MW;  F37E02F42C8717A2 CRC64;

  Query Match             5.1%;  Score 243.5;  DB 1;  Length 1178;
  Best Local Similarity   34.4%;  Pred. No. 1.7e-10;
  Matches   55;  Conservative   15;  Mismatches   75;  Indels   15;  Gaps    4;

Qy        210 QARLADTANYTCVAKNIVARRRS-ASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRS 268
              : |  |     |   :|  |  |        :  :||||  |: || || :||  |    | |
Db        410 RGRSCDVTRSACTGPHIQTRMCSFKKCDHRIRQDGGWSHWSPWSSCSVTCGVGNITRIRL 469

Qy        269 CTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDGSWSPWSKWSACGLDC---THWRSRECSD 325
              | :|  |  ||  | |                  |:|  |  ||| |||| :  |      ||| |:
Db        470 CNSPIPQMGGKNCVGNGRETEKCEKAPCPVNGQWGPWSPWSACTVTCGGGIRERSRLCNS 529

Qy        326 PAPRNGGEECQGTDLDT--------RNCTSDLCVHSASGP 357
              | |: ||: | |  ||         |:|  | |: :   |
Db        530 PEPQYGGKPCVG---DTKQHDMCNKRDCPIDGCLSNPCFP 566

RESULT 15
TSP1_BOVIN
ID   TSP1_BOVIN     STANDARD;      PRT;  1170 AA.
AC   Q28178; Q28179;
DT   01-NOV-1997 (Rel. 35, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Thrombospondin 1 precursor.
GN   THBS1 OR TSP1 OR TSP-1.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Holstein; TISSUE=Tooth;
RX   MEDLINE=98173773; PubMed=9507054;
RA   Ueno A., Yamashita K., Nagata T., Tsurumi C., Miwa Y., Kitamura S.,
RA   Inoue H.;
RT   "cDNA cloning of bovine thrombospondin 1 and its expression in
RT   odontoblasts and predentin.";
RL   Biochim. Biophys. Acta 1382:17-22(1998).
RN   [2]
RP   SEQUENCE OF 1-18 AND 710-1170 FROM N.A.
RC   TISSUE=Aortic endothelium;
RA   Zafar R.S., Moll Y.D., Womack J.F., Walz D.A.;
RL   Submitted (JUN-1995) to the EMBL/GenBank/DDBJ databases.
```

```
CC   -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC       cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC       laminin, type V collagen and integrins alpha-V/beta-1, alpha-
CC       V/beta-3 and alpha-IIb/beta-3. May play a role in dentinogenesis
CC       and/or maintenance of dentin and dental pulp.
CC   -!- SUBUNIT: Homotrimer; disulfide-linked.
CC   -!- TISSUE SPECIFICITY: Odontoblasts.
CC   -!- SIMILARITY: Belongs to the thrombospondin family.
CC   -!- SIMILARITY: Contains 1 VWFC domain.
CC   -!- SIMILARITY: Contains 3 EGF-like domains.
CC   -!- SIMILARITY: Contains 3 TSP type-1 domains.
CC   -!- SIMILARITY: Contains 7 TSP type-3 domains.
CC   -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AB005287; BAA21115.1; -.
DR   EMBL; X87618; CAA60950.1; -.
DR   EMBL; X87619; CAA60951.1; -.
DR   PIR; S55501; S55501.
DR   GlycoSuiteDB; Q28178; -.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR000884; TSP1.
DR   InterPro; IPR008085; TSP_1.
DR   InterPro; IPR003367; tsp_3.
DR   InterPro; IPR008859; TSPC.
DR   InterPro; IPR003129; TSPN.
DR   InterPro; IPR001007; VWF_C.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00090; tsp_1; 3.
DR   Pfam; PF02412; tsp_3; 13.
DR   Pfam; PF05735; TSPC; 1.
DR   Pfam; PF02210; TSPN; 1.
DR   Pfam; PF00093; vwc; 1.
DR   PRINTS; PR01705; TSP1REPEAT.
DR   SMART; SM00181; EGF; 3.
DR   SMART; SM00209; TSP1; 3.
DR   SMART; SM00210; TSPN; 1.
DR   SMART; SM00214; VWC; 1.
DR   PROSITE; PS00022; EGF_1; FALSE_NEG.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS50026; EGF_3; 2.
DR   PROSITE; PS50092; TSP1; 3.
DR   PROSITE; PS01208; VWFC_1; 1.
DR   PROSITE; PS50184; VWFC_2; 1.
KW   Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat;
KW   EGF-like domain; Signal.
FT   SIGNAL        1     18       BY SIMILARITY.
FT   CHAIN        19   1170       THROMBOSPONDIN 1.
FT   DOMAIN       19    232       HEPARIN-BINDING (POTENTIAL).
FT   DOMAIN       24    221       TSP N-TERMINAL.
FT   DOMAIN      316    373       VWFC.
FT   DOMAIN      379    429       TSP TYPE-1 1.
FT   DOMAIN      435    490       TSP TYPE-1 2.
FT   DOMAIN      492    547       TSP TYPE-1 3.
FT   DOMAIN      549    587       EGF-LIKE 1.
FT   DOMAIN      588    645       EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      646    690       EGF-LIKE 3.
FT   DOMAIN      723    758       TSP TYPE-3 1.
FT   DOMAIN      759    781       TSP TYPE-3 2.
FT   DOMAIN      782    817       TSP TYPE-3 3.
FT   DOMAIN      818    840       TSP TYPE-3 4.
FT   DOMAIN      841    878       TSP TYPE-3 5.
FT   DOMAIN      879    914       TSP TYPE-3 6.
FT   DOMAIN      915    950       TSP TYPE-3 7.
FT   DOMAIN      951   1170       C-TERMINAL.
FT   SITE        926    928       CELL ATTACHMENT SITE (POTENTIAL).
FT   DISULFID    270    270       INTERCHAIN (PROBABLE).
FT   DISULFID    274    274       INTERCHAIN (PROBABLE).
FT   DISULFID    391    423       BY SIMILARITY.
FT   DISULFID    395    428       BY SIMILARITY.
FT   DISULFID    406    413       BY SIMILARITY.
FT   DISULFID    447    484       BY SIMILARITY.
FT   DISULFID    451    489       BY SIMILARITY.
FT   DISULFID    462    474       BY SIMILARITY.
FT   DISULFID    504    541       BY SIMILARITY.
FT   DISULFID    508    546       BY SIMILARITY.
FT   DISULFID    519    531       BY SIMILARITY.
FT   DISULFID    551    562       BY SIMILARITY.
FT   DISULFID    556    572       BY SIMILARITY.
FT   DISULFID    575    586       BY SIMILARITY.
FT   DISULFID    592    608       BY SIMILARITY.
FT   DISULFID    599    617       BY SIMILARITY.
FT   DISULFID    620    644       BY SIMILARITY.
FT   DISULFID    650    663       BY SIMILARITY.
FT   DISULFID    657    676       BY SIMILARITY.
FT   DISULFID    678    689       BY SIMILARITY.
FT   DISULFID    705    713       BY SIMILARITY.
FT   DISULFID    718    738       BY SIMILARITY.
FT   DISULFID    754    774       BY SIMILARITY.
FT   DISULFID    777    797       BY SIMILARITY.
FT   DISULFID    813    833       BY SIMILARITY.
FT   DISULFID    836    856       BY SIMILARITY.
FT   DISULFID    874    894       BY SIMILARITY.
FT   DISULFID    910    930       BY SIMILARITY.
FT   DISULFID    946   1167       BY SIMILARITY.
FT   CARBOHYD    248    248       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    360    360       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    708    708       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1067   1067       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1085   1085       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    805    805       S -> G (IN REF. 2).
SQ   SEQUENCE   1170 AA;  129533 MW;  0DD6ADF3E5FA031A CRC64;

  Query Match             5.1%;  Score 243;  DB 1;  Length 1170;
  Best Local Similarity   30.3%;  Pred. No. 1.8e-10;
  Matches   47;  Conservative   24;  Mismatches   66;  Indels   18;  Gaps    3;

Qy        207 VVRQARLADTANYTCVAKNIVAR--------RRSASAAVIVYVNGGWSTWTEWSVCSASC 258
              : ::  |   |:  |   |   ::   |            :|          :||||  |:  ||  || :|
Db        399 IQQRGRSCDSLNNRCEGSSVQTRTCHIQECDKRFKQ-------DGGWSHWSPWSSCSVTC 451

Qy        259 GRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDGSWSPWSKWSACGLDC--- 315
              | |     |  |  |  :|:|     |   |||:       :          ::|  |  |||  |    |  :  |
Db        452 GDGVITRIRLCNSPSPQMNGKPCEGKARETKACQKDSCPINGGWGPWSPWDICSVTCGGG 511

Qy        316 THWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLC 350
                  |||  |::|  |:  ||::|  |     :  :  |        |
Db        512 VQKRSRLCNNPKPQFGGKDCVGDVTENQICNKQDC 546

Search completed: October  6, 2004, 18:18:00
Job time : 22 secs
```

```
OM protein - protein search, using sw model

Run on:          October  6, 2004, 17:50:33 ; Search time 84 Seconds
                                             (without alignments)
                                             3376.796 Million cell updates/sec

Title:           US-09-970-944-2
Perfect score:   4787
Sequence:        1 MAVRPGLWPALLGIVLAAWL..........AVAGLGQPDAGLFTVSEAEC 899

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters:      1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SPTREMBL_25:*
               1:  sp_archea:*
               2:  sp_bacteria:*
               3:  sp_fungi:*
               4:  sp_human:*
               5:  sp_invertebrate:*
               6:  sp_mammal:*
               7:  sp_mhc:*
               8:  sp_organelle:*
               9:  sp_phage:*
               10:  sp_plant:*
               11:  sp_rodent:*
               12:  sp_virus:*
               13:  sp_vertebrate:*
               14:  sp_unclassified:*
               15:  sp_rvirus:*
               16:  sp_bacteriap:*
               17:  sp_archeap:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 4592.5 | 95.9 | 898 | 11 | Q8K1S4 | Q8k1s4 mus musculu |
| 2 | 4545.5 | 95.0 | 898 | 11 | O08721 | O08721 rattus norv |
| 3 | 2819 | 58.9 | 544 | 4 | Q96GP4 | Q96gp4 homo sapien |
| 4 | 2705.5 | 56.5 | 931 | 11 | O08747 | O08747 mus musculu |
| 5 | 2686 | 56.1 | 950 | 11 | Q8CD16 | Q8cd16 mus musculu |
| 6 | 2679.5 | 56.0 | 931 | 13 | Q7T2Z5 | Q7t2z5 gallus gall |
| 7 | 2673.5 | 55.8 | 931 | 4 | O95185 | O95185 homo sapien |
| 8 | 2570 | 53.7 | 943 | 13 | Q8JGT4 | Q8jgt4 xenopus lae |
| 9 | 2505.5 | 52.3 | 1008 | 11 | Q80Y85 | Q80y85 mus musculu |
| 10 | 2499 | 52.2 | 945 | 11 | Q8K1S3 | Q8k1s3 mus musculu |
| 11 | 2497 | 52.2 | 945 | 11 | O08722 | O08722 rattus norv |
| 12 | 2493 | 52.1 | 945 | 11 | Q9D398 | Q9d398 mus musculu |
| 13 | 2486.5 | 51.9 | 934 | 4 | Q8IZJ1 | Q8izj1 homo sapien |
| 14 | 2479 | 51.8 | 945 | 4 | Q86SN3 | Q86sn3 homo sapien |
| 15 | 2144.5 | 44.8 | 956 | 11 | Q8K1S2 | Q8k1s2 mus musculu |
| 16 | 2136 | 44.6 | 948 | 4 | Q8WYP7 | Q8wyp7 homo sapien |
| 17 | 1597.5 | 33.4 | 597 | 4 | Q8IUT0 | Q8iut0 homo sapien |
| 18 | 1442 | 30.1 | 328 | 11 | Q80T71 | Q80t71 mus musculu |
| 19 | 1223.5 | 25.6 | 554 | 4 | Q8N1Y2 | Q8n1y2 homo sapien |
| 20 | 971 | 20.3 | 366 | 4 | Q9H9F3 | Q9h9f3 homo sapien |
| 21 | 961.5 | 20.1 | 1072 | 5 | Q9NBL0 | Q9nbl0 drosophila |
| 22 | 956.5 | 20.0 | 1072 | 5 | Q9V7B5 | Q9v7b5 drosophila |
| 23 | 953.5 | 19.9 | 947 | 5 | Q26262 | Q26262 caenorhabdi |
| 24 | 950.5 | 19.9 | 947 | 5 | O44171 | O44171 caenorhabdi |
| 25 | 694 | 14.5 | 199 | 13 | Q9PVD5 | Q9pvd5 petromyzon |
| 26 | 552.5 | 11.5 | 351 | 4 | Q8TF26 | Q8tf26 homo sapien |
| 27 | 366 | 7.6 | 2673 | 4 | Q96SC3 | Q96sc3 homo sapien |
| 28 | 366 | 7.6 | 5636 | 4 | Q96RW7 | Q96rw7 homo sapien |
| 29 | 311.5 | 6.5 | 325 | 5 | Q8I1K1 | Q8i1k1 drosophila |
| 30 | 306 | 6.4 | 518 | 4 | Q8IV45 | Q8iv45 homo sapien |
| 31 | 273.5 | 5.7 | 1172 | 11 | Q8CG21 | Q8cg21 mus musculu |
| 32 | 273.5 | 5.7 | 1172 | 11 | Q7TMT3 | Q7tmt3 mus musculu |
| 33 | 271 | 5.7 | 1582 | 11 | Q8CGM0 | Q8cgm0 mus musculu |
| 34 | 270.5 | 5.7 | 1081 | 5 | Q9U631 | Q9u631 drosophila |
| 35 | 269.5 | 5.6 | 1083 | 5 | Q9VTT0 | Q9vtt0 drosophila |
| 36 | 269.5 | 5.6 | 1091 | 5 | Q7YU67 | Q7yu67 drosophila |
| 37 | 267 | 5.6 | 1522 | 11 | Q80ZF8 | Q80zf8 mus musculu |
| 38 | 263.5 | 5.5 | 1461 | 5 | Q8MYA8 | Q8mya8 caenorhabdi |
| 39 | 258.5 | 5.4 | 1122 | 11 | Q7TT33 | Q7tt33 mus musculu |
| 40 | 258 | 5.4 | 1573 | 4 | Q8NGW8 | Q8ngw8 homo sapien |
| 41 | 256 | 5.3 | 478 | 11 | Q8BVE5 | Q8bve5 mus musculu |
| 42 | 254.5 | 5.3 | 1560 | 11 | Q8CGM1 | Q8cgm1 mus musculu |
| 43 | 249 | 5.2 | 685 | 6 | Q9TTS5 | Q9tts5 bos taurus |
| 44 | 249 | 5.2 | 5146 | 6 | Q8SPM4 | Q8spm4 bos taurus |
| 45 | 248 | 5.2 | 1171 | 11 | Q8CGB2 | Q8cgb2 mus musculu |

ALIGNMENTS

```
RESULT 1
Q8K1S4
ID   Q8K1S4      PRELIMINARY;      PRT;   898 AA.
AC   Q8K1S4;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Netrin receptor Unc5h1.
GN   UNC5H1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Engelkamp D.;
RT   "Cloning of three mouse unc-5 genes and their expression patterns at
RT   mid-gestation.";
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ487852; CAD32250.1; -.
DR   MGD; MGI:894682; Unc5h1.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0007165; P:signal transduction; IEA.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR000884; TSP1.
DR   InterPro; IPR008085; TSP_1.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00090; tsp_1; 2.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01705; TSP1REPEAT.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00409; IG; 1.
DR   SMART; SM00209; TSP1; 2.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
```

```
DR   PROSITE; PS50092; TSP1; 2.
KW   Receptor.
SQ   SEQUENCE   898 AA;  98856 MW;  59F04BA2E196C1DB CRC64;

  Query Match             95.9%;  Score 4592.5;  DB 11;  Length 898;
  Best Local Similarity   95.3%;  Pred. No. 0;
  Matches  858;  Conservative   21;  Mismatches   18;  Indels    3;  Gaps    3;

Qy          1 MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60
Db          1 MAVRPGLWPALLGIVLTAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60

Qy         61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGEPTMEVRINVSRQQVEKVFGLE 120
Db         61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSRQQVEKVFGLE 120

Qy        121 EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180
Db        121 EYWCQCVAWSSSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180

Qy        181 VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY 240
Db        181 VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY 240

Qy        241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDG 300
Db        241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNV-QKTACATLCPVDG 299

Qy        301 SWSPWSKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCVHSASGPEDV 360
Db        300 SWSPWSKWSACGLDCTHWRSRECSDPAPRNGGEECRGADLDTRNCTSDLCLHTSSGPEDV 359

Qy        361 ALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHL 420
Db        360 ALYIGLVAVAVCLILLLLVLVLIYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHL 419

Qy        421 LTIQPDLS-TTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEEFV 479
Db        420 LTIQPDLSTTTTTYQGSLCPRQDGPSPKFQLSNGHLLSPLGSGRHTLHHSSPTSEAEDFV 479

Qy        480 SRLSTQNYFRSLPRGTSNMTYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLH 539
Db        480 SRLSTQNYFRSLPRGTSNMAYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLH 539

Qy        540 KPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGS 599
Db        540 KPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGS 599

Qy        600 WEQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAAKRLKLLLFAPVA 659
Db        600 WE-DVLHLGEESPSHLYYCQLEAGACYVFTEQLGRFALVGEALSVAATKRLRLLLFAPVA 658

Qy        660 CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSS 719
Db        659 CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSS 718

Qy        720 LWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSI 779
Db        719 LWKSKLLVSYQEIPFYHIWNGTQQYLHCTFTLERVNASTSDLACKVWVWQVEGDGQSFNI 778

Qy        780 NFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQ 839
Db        779 NFNITKDTRFAEMLALESEGGVPALVGPSAFKIPFLIRQKIITSLDPPCSRGADWRTLAQ 838

Qy        840 KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 899
Db        839 KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLGQLAAAVAGLGQPDAGLFTVSEAEC 898
```

```
RESULT 2
O08721
ID   O08721      PRELIMINARY;      PRT;   898 AA.
AC   O08721;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Transmembrane receptor UNC5H1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain, and Ventral spinal cord;
RX   MEDLINE=97271897; PubMed=9126742;
RA   Leonardo E.D., Hinck L., Masu M., Keino-Masu K., Ackerman S.L.,
RA   Tessier-Lavigne M.;
RT   "Vertebrate homologues of C. elegans UNC-5 are candidate netrin
RT   receptors.";
RL   Nature 386:833-838(1997).
DR   EMBL; U87305; AAB57678.1; -.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0007165; P:signal transduction; IEA.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR000884; TSP1.
DR   InterPro; IPR008085; TSP_1.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00090; tsp_1; 2.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01705; TSP1REPEAT.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00409; IG; 1.
DR   SMART; SM00209; TSP1; 2.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS50092; TSP1; 2.
KW   Receptor.
SQ   SEQUENCE   898 AA;  98840 MW;  7A3CBCB9E7ACA135 CRC64;

  Query Match             95.0%;  Score 4545.5;  DB 11;  Length 898;
  Best Local Similarity   94.7%;  Pred. No. 0;
  Matches  852;  Conservative   19;  Mismatches   26;  Indels    3;  Gaps    3;

Qy          1 MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60
Db          1 MAVRPGLWPVLLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60

Qy         61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGEPTMEVRINVSRQQVEKVFGLE 120
Db         61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDSSSGLPTMEVRINVSRQQVEKVFGLE 120

Qy        121 EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180
Db        121 EYWCQCVAWSSSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180

Qy        181 VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY 240
Db        181 VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSTSAAVIVY 240

Qy        241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDG 300
Db        241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNV-QKTACATLCPVDG 299

Qy        301 SWSPWSKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCVHSASGPEDV 360
Db        300 SWSSWSKWSACGLDCTHWRSRECSDPAPRNGGEECRGADLDTRNCTSDLCLHTASCPEDV 359

Qy        361 ALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHL 420
Db        360 ALYIGLVAVAVCLFLLLLALGLIYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHL 419

Qy        421 LTIQPDLS-TTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEEFV 479
```

```
Db        420 LTIQPDLSTTTTTYQGSLCSRQDGPSPKFQLSNGHLLSPLGSGRHTLHHSSPTSEAEDFV 479

Qy        480 SRLSTQNYFRSLPRGTSNMTYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLH 539
Db        480 SRLSTQNYFRSLPRGTSNMAYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLH 539

Qy        540 KPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGS 599
Db        540 KPEDVRLPLAGCQTLLSPVVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGS 599

Qy        600 WEQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAAKRLKLLLFAPVA 659
Db        600 WE-DVLHLGEESPSHLYYCQLEAGACYVFTEQLGRFALVGEALSVAATKRLRLLLFAPVA 658

Qy        660 CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSS 719
Db        659 CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSS 718

Qy        720 LWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSI 779
Db        719 LWKSKLLVSYQEIPFYHIWNGTQQYLHCTFTLERINASTSDLACKVWVWQVEGDGQSFNI 778

Qy        780 NFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQ 839
Db        779 NFNITKDTRFAELLALESEGGVPALVGPSAFKIPFLIRQKIIASLDPPCSRGADWRTLAQ 838

Qy        840 KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 899
Db        839 KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLGQLAAAVAGLGQPDAGLFTVSEAEC 898
```

RESULT 3
Q96GP4

```
ID   Q96GP4      PRELIMINARY;      PRT;   544 AA.
AC   Q96GP4;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Similar to transmembrane receptor Unc5H1 (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RA   Strausberg R.;
RL   Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC009333; AAH09333.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0007165; P:signal transduction; IEA.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00791; ZU5; 1.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00218; ZU5; 1.
KW   Receptor; Transmembrane.
FT   NON_TER       1      1
SQ   SEQUENCE   544 AA;  59949 MW;  350A7BA53375CCAE CRC64;

  Query Match             58.9%;  Score 2819;  DB 4;  Length 544;
  Best Local Similarity   99.6%;  Pred. No. 8.2e-252;
  Matches  540;  Conservative    0;  Mismatches    0;  Indels    2;  Gaps    2;

Qy        359 DVALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNP 418
Db          4 DVALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNP 63

Qy        419 HLLTIQPDLS-TTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEE 477
Db         64 HLLTIQPDLSTTTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEE 123

Qy        478 FVSRLSTQNYFRSLPRGTSNMTYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLT 537
Db        124 FVSRLSTQNYFRSLPRGTSNMTYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLT 183

Qy        538 LHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCE 597
Db        184 LHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCE 243

Qy        598 GSWEQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAAKRLKLLLFAP 657
Db        244 GSWE-DVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAAKRLKLLLFAP 302

Qy        658 VACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVP 717
Db        303 VACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVP 362

Qy        718 SSLWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSF 777
Db        363 SSLWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSF 422

Qy        778 SINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTL 837
Db        423 SINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTL 482

Qy        838 AQKLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEA 897
Db        483 AQKLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEA 542

Qy        898 EC 899
Db        543 EC 544
```

RESULT 4
O08747

```
ID   O08747      PRELIMINARY;      PRT;   931 AA.
AC   O08747;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Rostral cerebellar malformation protein.
GN   UNC5H3 OR RCM.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57B6/SJL;
RX   MEDLINE=97271898; PubMed=9126743;
RA   Ackerman S.L., Kozak L.P., Przyborski S.A., Rund L.A., Boyer B.B.,
RA   Knowles B.B.;
RT   "The mouse rostral cerebellar malformation gene encodes an UNC-5-like
RT   protein.";
RL   Nature 386:838-842(1997).
DR   EMBL; U72634; AAB54103.1; -.
DR   MGD; MGI:1095412; Unc5h3.
DR   GO; GO:0005886; C:plasma membrane; IC.
DR   GO; GO:0005042; F:netrin receptor activity; IDA.
DR   GO; GO:0005515; F:protein binding; IDA.
DR   GO; GO:0007420; P:brain development; IMP.
DR   GO; GO:0030334; P:regulation of cell migration; IMP.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR000884; TSP1.
DR   InterPro; IPR008085; TSP_1.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00090; tsp_1; 2.
```

```
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01705; TSP1REPEAT.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00408; IGc2; 1.
DR   SMART; SM00209; TSP1; 2.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS50092; TSP1; 2.
KW   Immunoglobulin domain.
SQ   SEQUENCE   931 AA;  103062 MW;  8A5D951A4EECA179 CRC64;

  Query Match             56.5%;  Score 2705.5;  DB 11;  Length 931;
  Best Local Similarity   56.4%;  Pred. No. 6.2e-241;
  Matches  515;  Conservative  153;  Mismatches  214;  Indels   31;  Gaps   12;

Qy          9 PALLGIVLAAWLRGSGAQQSA---TVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVC 65
Db         26 PAL--ALLSASGTGSAAQDDEFFHELPETFPSDPPEPLPHFLIEPEEAYIVKNKPVNLYC 83

Qy         66 KAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGEPTMEVRINVSRQQVEKVFGLEEYWCQ 125
Db         84 KASPATQIYFKCNSEWVHQKDHVVDERVDETSGLIVREVSIEISRQQVEELFGPEDYWCQ 143

Qy        126 CVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAEVEWLR 185
Db        144 CVAWSSAGTTKSRKAYVRIAYLRKTFEQEPLGKEVSLEQEVLLQCRPPEGIPVAEVEWLK 203

Qy        186 NEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNGGW 245
Db        204 NEDIIDPAEDRNFYITIDHNLIIKQARLSDTANYTCVAKNIVAKRKSTTATVIVYVNGGW 263

Qy        246 STWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDGSWSPW 305
Db        264 STWTEWSVCNSRCGRGYQKRTRTCTNPAPLNGGAFCEGQSV-QKIACTTLCPVDGRWTSW 322

Qy        306 SKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCVHSASGPEDVALYVG 365
Db        323 SKWSTCGTECTHWRRRECTAPAPKNGGKDCDGLVLQSKNCTDGLCMQAAPDSDDVALYVG 382

Qy        366 L-IAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHLLTIQ 424
Db        383 IVIAVTVCLAITVVVALFVYRKNHRDFESDIIDSSALNGGFQPVNIKAARQD---LLAVP 439

Qy        425 PDL-STTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSS----PTSEAEEFV 479
Db        440 PDLTSAAAMYRGPVYALHD-VSDKIPMTNSPILDPLPNLKIKVYNSSGAVTPQDDLAEFS 498

Qy        480 SRLS---TQNYF---------RSLPRGT--SNMTYGTFNFLGGRLMIPNTGISLLIPPDA 525
Db        499 SKLSPQMTQSLLENEALNLKNQSLARQTDPSCTAFGTFNSLGGHLIIPNSGVSLLIPAGA 558

Qy        526 IPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPD 585
Db        559 IPQGRVYEMYVTVHRKENMRPPMEDSQTLLTPVVSCGPPGALLTRPVILTLHHCADPSTE 618

Qy        586 SWSLRLKKQSCEGSWEQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA 645
Db        619 DWKIQLKNQAVQGQWE-DVVVVGEENFTTPCYIQLDAEACHILTENLSTYALVGQSTTKA 677

Qy        646 AAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDS 705
Db        678 AAKRLKLAIFGPLCCSSLEYSIRVYCLDDTQDALKEVLQLERQMGGQLLEEPKALHFKGS 737

Qy        706 YHNLRLSIHDVPSSLWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKL 765
Db        738 IHNLRLSIHDIAHSLWKSKLLAKYQEIPFYHIWSGSQRNLHCTFTLERLSLNTVELVCKL 797

Qy        766 WVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLD 825
Db        798 CVRQVEGEGQIFQLNCTVSEEPTGIDLPLLDPASTITTVTGPSAFSIPLPIRQKLCSSLD 857

Qy        826 PPCRRGADWRTLAQKLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLG 885
Db        858 APQTRGHDWRMLAHKLNLDRYLNYFATKSSPTGVILDLWEAQNFPDGNLSMLAAVLEEMG 917

Qy        886 QPDAGLFTVSEAE 898
Db        918 RHETVVSLAAEGQ 930


RESULT 5
Q8CD16
ID   Q8CD16      PRELIMINARY;      PRT;   950 AA.
AC   Q8CD16;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Unc5 homolog.
GN   UNC5H3.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Testis;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK031655; BAC27495.1; -.
DR   PIR; PT0566; PT0566.
DR   MGD; MGI:1095412; Unc5h3.
DR   GO; GO:0005886; C:plasma membrane; IC.
DR   GO; GO:0005042; F:netrin receptor activity; IDA.
DR   GO; GO:0005515; F:protein binding; IDA.
DR   GO; GO:0007420; P:brain development; IMP.
DR   GO; GO:0030334; P:regulation of cell migration; IMP.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR000884; TSP1.
DR   InterPro; IPR008085; TSP_1.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00090; tsp_1; 2.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01705; TSP1REPEAT.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00409; IG; 1.
DR   SMART; SM00408; IGc2; 1.
DR   SMART; SM00209; TSP1; 2.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS50092; TSP1; 2.
SQ   SEQUENCE   950 AA;  105398 MW;  1E8FC74703351AF6 CRC64;

  Query Match             56.1%;  Score 2686;  DB 11;  Length 950;
  Best Local Similarity   55.3%;  Pred. No. 4.1e-239;
  Matches  515;  Conservative  153;  Mismatches  214;  Indels   50;  Gaps   13;

Qy          9 PALLGIVLAAWLRGSGAQQSA---TVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVC 65
Db         26 PAL--ALLSASGTGSAAQDDEFFHELPETFPSDPPEPLPHFLIEPEEAYIVKNKPVNLYC 83

Qy         66 KAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGEPTMEVRINVSRQQVEKVFGLEEYWCQ 125
Db         84 KASPATQIYFKCNSEWVHQKDHVVDERVDETSGLIVREVSIEISRQQVEELFGPEDYWCQ 143

Qy        126 CVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAEVEWLR 185
Db        144 CVAWSSAGTTKSRKAYVRIAYLRKTFEQEPLGKEVSLEQEVLLQCRPPEGIPVAEVEWLK 203
```

```
Qy        186 NEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNGGW 245
Db        204 NEDIIDPAEDRNFYITIDHNLIIKQARLSDTANYTCVAKNIVAKRKSTTATVIVYVNGGW 263

Qy        246 STWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDGSWSPW 305
Db        264 STWTEWSVCNSRCGRGYQKRTRTCTNPAPLNGGAFCEGQSV-QKIACTTLCPVDGRWTSW 322

Qy        306 SKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCVH------------ 352
Db        323 SKWSTCGTECTHWRRRECTAPAPKNGGKDCDGLVLQSKNCTDGLCMQGFIYPISTEHRPQ 382

Qy        353 ------SASGPEDVALYVGL-IAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTSGF 405
Db        383 NEYGFSSAPDSDDVALYVGIVIAVTVCLAITVVVALFVYRKNHRDFESDIIDSSALNGGF 442

Qy        406 QPVSIKPSKADNPHLLTIQPDL-STTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRH 464
Db        443 QPVNIKAARQD---LLAVPPDLTSAAAMYRGPVYALHD-VSDKIPMTNSPILDPLPNLKI 498

Qy        465 TLHHSS----PTSEAEEFVSRLS---TQNYF---------RSLPRGT--SNMTYGTFNFL 506
Db        499 KVYNSSGAVTPQDDLAEFSSKLSPQMTQSLLENEALNLKNQSLARQTDPSCTAFGTFNSL 558

Qy        507 GGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGV 566
Db        559 GGHLIIPNSGVSLLIPAGAIPQGRVYEMYVTVHRKENMRPPMEDSQTLLTPVVSCGPPGA 618

Qy        567 LLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGSWEQDVLHLGEEAPSHLYYCQLEASACY 626
Db        619 LLTRPVILTLHHCADPSTEDWKIQLKNQAVQGQWE-DVVVVGEENFTTPCYIQLDAEACH 677

Qy        627 VFTEQLGRFALVGEALSVAAAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLE 686
Db        678 ILTENLSTYALVGQSTTKAAAKRLKLAIFGPLCCSSLEYSIRVYCLDDTQDALKEVLQLE 737

Qy        687 KQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWKSKLLVSYQEIPFYHIWNGTQRYLH 746
Db        738 RQMGGQLLEEPKALRFKGSIHNLRLSIHDIAHSLWKSKLLAKYQEIPFYHIWSGSQRNLH 797

Qy        747 CTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVG 806
Db        798 CTFTLERLSLNTVELVCKLCVRQVEGEGQIFQLNCTVSEEPTGIDLPLLDPASTITTVTG 857

Qy        807 PSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLDSHLSFFASKPSPTAMILNLWEA 866
Db        858 PSAFSIPLPIRQKLCSSLDAPQTRGHDWRMLAHKLNLDRYLNYFATKSSPTGVILDLWEA 917

Qy        867 RHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE 898
Db        918 QNFPDGNLSMLAAVLEEMGRHETVVYLAAEGQ 949
```

RESULT 6
Q7T2Z5

```
ID   Q7T2Z5      PRELIMINARY;      PRT;   931 AA.
AC   Q7T2Z5;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   UNC5-like protein 3.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Guan W., Condic M.L.;
RT   "Characterization of Netrin-1, Neogenin and cUNC-5H3 expression during
RT   chick dorsal root ganglia development.";
RL   Gene Expr. Patterns 3:369-373(2003).
DR   EMBL; AY187310; AAO67275.1; -.
SQ   SEQUENCE   931 AA;  102906 MW;  1E23A0D84F2E2C62 CRC64;

  Query Match             56.0%;  Score 2679.5;  DB 13;  Length 931;
  Best Local Similarity   56.1%;  Pred. No. 1.6e-238;
  Matches  511;  Conservative  151;  Mismatches  218;  Indels   31;  Gaps   12;

Qy          9 PALLGIVLAAWLRGSGAQQS---ATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVC 65
Db         26 PAL--AVLGASRPGSAAQDDDFFHELPETFPSDPPEPLPHFLIEPEEAYIVKNKPVNLYC 83

Qy         66 KAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGEPTMEVRINVSRQQVEKVFGLEEYWCQ 125
Db         84 KASPATQIYFKCNSEWVHQKDHVVDERVDETSGLIVCEVSIEISRQQVEELFGPEDYWCQ 143

Qy        126 CVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAEVEWLR 185
Db        144 CVAWSSAGTTKSRKAYVRIAYLRKTFEQEPLGKEVSLEQEVLLQCRPPEGIPVAEVEWLK 203

Qy        186 NEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNGGW 245
Db        204 NEEVIDPVEDRNFYITIDHNLIIKQARLSDTANYTCVAKNIVAKRKSTTATVIVYVNGGW 263

Qy        246 STWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDGSWSPW 305
Db        264 STWTEWSACNSRCGRGFQKRTRTCTNPAPLNGGAFCEGQNV-QKIACTTLCPVDGKWTSW 322

Qy        306 SKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCVHSASGPEDVALYVG 365
Db        323 SKWSTCGTECTHWRRRECTAPAPKNGGKDCEGLVLQSKNCTDGLCMQAAPDSDDVALYVG 382

Qy        366 L-IAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHLLTIQ 424
Db        383 IVIAVIVCLAISVVVALFVYRKNHRDFESDIIDSSALNGGFQPVNIKAARQD---LLAVP 439

Qy        425 PDL-STTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSS----PTSEAEEFV 479
Db        440 PDLTSAAAMYRGPVYALHD-VSDKIPMTNSPILDPLPNLKIKVYNTSGAVTPQDELSDFS 498

Qy        480 SRLS---TQNYF---------RSLPRGT--SNMTYGTFNFLGGRLMIPNTGISLLIPPDA 525
Db        499 SKLSPQITQSLLENETLNVKNQSLARQTDPSCTAFGTFNSLGGHLVIPNSGVSLLIPAGA 558

Qy        526 IPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPD 585
Db        559 VPQGRVYEMYVTVHRKEGMRPPVEDSQTLLTPVVSCGPPGALLTRPVVLTMHHCAEPNMD 618

Qy        586 SWSLRLKKQSCEGSWEQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA 645
Db        619 DWQIQLKHQAGQGPWE-DVVVVGEENFTTPCYIQLDPEACHILTETLSTYALVGQSITKA 677

Qy        646 AAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDS 705
Db        678 AAKRLKLAIFGPLSCSSLEYSIRVYCLDDTQDALKEVLQLERQMGGQLLEEPKTLHFKGS 737

Qy        706 YHNLRLSIHDVPSSLWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKL 765
Db        738 THNLRLSIHDIAHSLWKSKLPAKYQEIPFYHIWSGCQRNLHCTFTLERFSLNTLELVCKL 797

Qy        766 WVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLD 825
Db        798 CVRQVEGEGQIFQLNCSVSEEPTGIDYPIMDSAGSITTIVGPNAFSIPLPIRQKLCSSLD 857

Qy        826 PPCRRGADWRTLAQKLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLG 885
Db        858 APQTRGHDWRMLAHKLKLDRYLNYFATKSSPTGVILDLWEAQNFPDGNLSMLAAVLEEMG 917

Qy        886 QPDAGLFTVSE 896
Db        918 RHETVVSLAAE 928
```

RESULT 7

```
O95185
ID   O95185      PRELIMINARY;      PRT;   931 AA.
AC   O95185;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Transmembrane receptor UNC5C.
GN   UNC5C.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=99000841; PubMed=9782087;
RA   Ackerman S.L., Knowles B.B.;
RT   "Cloning and mapping of the UNC5C gene to human chromosome 4q21-q23.";
RL   Genomics 52:205-208(1998).
DR   EMBL; AF055634; AAC67491.1; -.
DR   Genew; HGNC:12569; UNC5C.
DR   GO; GO:0005042; F:netrin receptor activity; TAS.
DR   GO; GO:0007411; P:axon guidance; TAS.
DR   GO; GO:0007420; P:brain development; TAS.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR000884; TSP1.
DR   InterPro; IPR008085; TSP_1.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00090; tsp_1; 2.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01705; TSP1REPEAT.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00408; IGc2; 1.
DR   SMART; SM00209; TSP1; 2.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS50092; TSP1; 2.
KW   Immunoglobulin domain; Receptor.
SQ   SEQUENCE   931 AA;  103101 MW;  EFD71122C98DABB8 CRC64;

  Query Match             55.8%;  Score 2673.5;  DB 4;  Length 931;
  Best Local Similarity   55.5%;  Pred. No. 5.7e-238;
  Matches  507;  Conservative  154;  Mismatches  221;  Indels   31;  Gaps   12;

Qy          9 PALLGIVLAAWLRGSGAQQS---ATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVC 65
Db         26 PAL--ALLSASGTGSAAQDDDFFHELPETFPSDPPEPLPHFLIEPEEAYIVKNKPVNLYC 83

Qy         66 KAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGEPTMEVRINVSRQQVEKVFGLEEYWCQ 125
Db         84 KASPATQIYFKCNSEWVHQKDHIVDERVDETSGLIVREVSIEISRQQVEELFGPEDYWCQ 143

Qy        126 CVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAEVEWLR 185
Db        144 CVAWSSAGTTKSRKAYVRIAYLRKTFEQEPLGKEVSLEQEVLLQCRPPEGIPVAEVEWLK 203

Qy        186 NEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNGGW 245
Db        204 NEDIIDPVEDRNFYITIDHNLIIKQARLSDTANYTCVAKNIVAKRKSTTATVIVYVNGGW 263

Qy        246 STWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDGSWSPW 305
Db        264 STWTEWSVCNSRCGRGYQKRTRTCTNPAPLNGGAFCEGQSV-QKIACTTLCPVDGRWTPW 322

Qy        306 SKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCVHSASGPEDVALYVG 365
Db        323 SKWSTCGTECTHWRRRECTAPAPKNGGKDCDGLVLQSKNCTDGLCMQTAPDSDDVALYVG 382

Qy        366 L-IAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHLLTIQ 424
Db        383 IVIAVIVCLAISVVVALFVYRKNHRDFESDIIDSSALNGGFQPVNIKAARQD---LLAVP 439

Qy        425 PDL-STTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHS----SPTSEAEEFV 479
Db        440 PDLTSAAAMYRGPVYALHD-VSDKIPMTNSPILDPLPNLKIKVYNTSGAVSPQDDLSEFT 498

Qy        480 SRLS---TQNYF---------RSLPRGT--SNMTYGTFNFLGGRLMIPNTGISLLIPPDA 525
Db        499 SKLSPQMTQSLLENEALSLKNQSLARQTDPSCTAFGSFNSLGGHLIVPNSGVSLLIPAGA 558

Qy        526 IPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPD 585
Db        559 IPQGRVYEMYVTVHRKETMRPPMDDSQTLLTPVVSCGPPGALLTRPVVLTMHHCADPNTE 618

Qy        586 SWSLRLKKQSCEGSWEQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA 645
Db        619 DWKILLKNQAAQGQWE-DVVVVGEENFTTPCYIKLDAEACHILTENLSTYALVGHSTTKA 677

Qy        646 AAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDS 705
Db        678 AAKRLKLAIFGPLCCSSLEYSIRVYCLDDTQDALKEILHLERQTGGQLLEEPKALHFKGS 737

Qy        706 YHNLRLSIHDVPSSLWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKL 765
Db        738 THNLRLSIHDIAHSLWKSKLLAKYQEIPFYHVWSGSQRNLHCTFTLERFSLNTVELVCKL 797

Qy        766 WVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLD 825
Db        798 CVRQVEGEGQIFQLNCTVSEEPTGIDLPLLDPANTITTVTGPSAFSIPLPIRQKLCSSLD 857

Qy        826 PPCRRGADWRTLAQKLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLG 885
Db        858 APQTRGHDWRMLAHKLNLDRYLNYFATKSSPTGVILDLWEAQNFPDGNLSMLAAVLEEMG 917

Qy        886 QPDAGLFTVSEAE 898
Db        918 RHETVVSLAAEGQ 930

RESULT 8
Q8JGT4
ID   Q8JGT4      PRELIMINARY;      PRT;   943 AA.
AC   Q8JGT4;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   UNC-5 receptor.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Anderson R.B., Holt C.E.;
RT   "Expression of UNC-5 in the developing Xenopus visual system.";
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY099459; AAM34486.1; -.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0007165; P:signal transduction; IEA.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR000884; TSP1.
DR   InterPro; IPR008085; TSP_1.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00090; tsp_1; 2.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01705; TSP1REPEAT.
```

```
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00409; IG; 2.
DR   SMART; SM00408; IGc2; 1.
DR   SMART; SM00209; TSP1; 2.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS50092; TSP1; 2.
KW   Immunoglobulin domain; Receptor.
SQ   SEQUENCE   943 AA;  105083 MW;  A024E24A7EDB6175 CRC64;

  Query Match             53.7%;  Score 2570;  DB 13;  Length 943;
  Best Local Similarity   52.1%;  Pred. No. 2.2e-228;
  Matches  488;  Conservative  166;  Mismatches  233;  Indels   50;  Gaps   11;

Qy         10 ALLGIVLAAWLRG-------SGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVL 62
              | |  :| | :          :| : |  : :  | |  : |||||:|||| ||||||||
Db         10 AALAAILVALILSCNFPSSTAGIEYSDVLPDSFPSAPAESLPHFLLEPEDAYIVKNKPVE 69

Qy         63 LVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGEPTMEVRINVSRQQVEKVFGLEEY 122
              |||| |||||:|||||||| | ||: :    |  :|      ||:| |||||||::||||:|
Db         70 LVCKANPATQIYFKCNGEWVNQNDHITKERVDDVTGLVVREVQIEVSRQQVEELFGLEDY 129

Qy        123 WCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAEVE 182
              ||||||||:|||||:::|:||| |||||:|||| |||:|||   :| ||||||:||||||
Db        130 WCQCVAWSSAGTTKSKRSYVRIAYLRKNFDQEPLGKEVALEQEALLQCRPPEGVPPAEVE 189

Qy        183 WLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVN 242
              ||:|:::||: |  | || :|:|:::||||:|||||||:||||||:||| :| |||:||
Db        190 WLKNEEIIDPTKDTNFLITIDHNLIIKQARLSDTANYTCVSKNIVAKRRSTTATVIVFVN 249

Qy        243 GGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDGSW 302
              ||||:||||| |:  || |||||:|:||||||||||  ||||   :     :   ||| |
Db        250 GGWSSWTEWSPCNNRCGHGWQKRTRTCTNPAPLNGGTMCEGQQ-YQKFACNTMCPVDGGW 308

Qy        303 SPWSKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCVHS--------- 353
              : |||||| :||||||||: | |:|||:| |  ||::|||  ||: :
Db        309 TEWSKWSACSTECTHWRSRECNAPTPKNGGKDCSGMLLDSKNCTDGLCMQNKRVLGETKS 368

Qy        354 --ASGPEDVALYVGL-IAVAVCLVLLLLVLILVYCRKKEGLDSDVADSS-ILTSGFQPVS 409
                      ||||| || :|: : ::||: | |:|| |      |:|: |||  || || ||:
Db        369 RLLESTGDVALYAGLVVAIFIVIILLMAVGIVVYRRNCREFDTDITDSSAALTGGFHPVN 428

Qy        410 IKPSKADNPHLL--TIQPDLSTTT-TYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTL 466
               | |: ||  |:   :||||:       |:|::    ||  :  |  :||  || ||    :  :
Db        429 FKTSRHDNSQLIHPAMQPDLTANAGIYRGNMYALQDS-ADKIPMTNSPLLDPLPNLKIKV 487

Qy        467 HHSS------------------PTSEAEEFVSRLSTQN------YFRSLPRGTSNMTYGT 502
              ::||                  ||         : :: :|          :|||  :||  ||
Db        488 YNSSTVGSSPGIHDGNNLLGTKPTGTYPSDNNIMNARNKNMSMQHLLTLPRDSSNSVTGT 547

Qy        503 FNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCG 562
              |  |||||  ||||:|||||  |||:|| ||:|| ::| |:  ||   | ||:||||::||
Db        548 FGSLGGRLTFPNTGVSLLIPQGAIPQGKYYEMYLMINKRENTVLPSEGTQTILSPIITCG 607

Qy        563 PPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGSWEQDVLHLGEEAPSHLYYCQLEA 622
              | |:|| :|||  : ||  : :     | |:|| || :|:|| :|: | ||   :    ||||:
Db        608 PTGLLLCKPVILTVPHCADINTSDWILQLKTQSHQGNWE-EVVTLNEETLNTPCYCQLES 666

Qy        623 SACYVFTEQLGRFALVGEALSVAAAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEV 682
               :|:    :||| :| |||: | :| |||:| :|||: |||||||::|||: || ||||||
Db        667 HSCHTLLDQLGTYAFVGESYSRSAIKRLQLAIFAPMLCTSLEYNLKVYCVEDTPDALKEV 726

Qy        683 VQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWKSKLLVSYQEIPFYHIWNGTQ 742
              ::||| ||| |::||::| ||||||||||||||:| |||:|||:    ||||||||||:|:|
Db        727 LELEKTLGGYLVEEPKLLMFKDSYHNLRLSIHDIPHSLWRSKLMAKYQEIPFYHIWSGSQ 786

Qy        743 RYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVP 802
              | ||||||||| | : ::| ||: | ||||:|| | ::  : :: : :   : :  ::| |
Db        787 RTLHCTFTLERYSLAATELTCKICVRQVEGEGQIFQLHTLLEENVKSFDPFCSQAENSVT 846

Qy        803 ALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLDSHLSFFASKPSPTAMILN 862
               :|| |||||| |||| :||| |  || ||| ||||| :| :|::||:| ||| :||:
Db        847 THLGPYAFKIPFSIRQKICNSLDAPNSRGNDWRLLAQKLCMDRYLNYFATKASPTGVILD 906

Qy        863 LWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 899
              |||| |  :|:|: ||:|:  :|: :  |   :: :|
Db        907 LWEALHQDDGDLNTLASALEEMGKSEMMLVMATDGDC 943

RESULT 9
Q80Y85
ID   Q80Y85      PRELIMINARY;      PRT;  1008 AA.
AC   Q80Y85;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Unc5h2 protein (Fragment).
GN   UNC5H2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6; TISSUE=Brain;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6; TISSUE=Brain;
RA   Strausberg R.;
RL   Submitted (MAR-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC048162; AAH48162.1; -.
DR   GO; GO:0007165; P:signal transduction; IEA.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR000884; TSP1.
DR   InterPro; IPR008085; TSP_1.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00090; tsp_1; 2.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01705; TSP1REPEAT.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00409; IG; 2.
DR   SMART; SM00408; IGc2; 1.
DR   SMART; SM00209; TSP1; 2.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS50092; TSP1; 2.
FT   NON_TER       1      1
SQ   SEQUENCE   1008 AA;  110438 MW;  BCE5CA0EC537C130 CRC64;
```

```
  Query Match             52.3%;  Score 2505.5;  DB 11;  Length 1008;
  Best Local Similarity   52.7%;  Pred. No. 2.3e-222;
  Matches  497;  Conservative  153;  Mismatches  240;  Indels   53;  Gaps   17;

Qy          1 MAVRPGLWPALLGIVLAAW-----LRG--SGAQQSATVANPVPGANPDLLPHFLVEPEDV 53
Db         75 MRARSGVRSALLLALLLCWDPTPSLAGVDSAGQ---VLPDSYPSAPAEQLPYFLLEPQDA 131

Qy         54 YIVKNKPVLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGEPTMEVRINVSRQQV 113
Db        132 YIVKNKPVELHCRAFPATQIYFKCNGEWVSQNDHVTQESLDEATGLRVREVQIEVSRQQV 191

Qy        114 EKVFGLEEYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPP 173
Db        192 EELFGLEDYWCQCVAWSSSGTTKSRRAYIRIAYLRKNFDQEPLAKEVPLDHEVLLQCRPP 251

Qy        174 EGIPPAEVEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSA 233
Db        252 EGVPVAEVEWLKNEDVIDPAQDTNFLLTIDHNLIIRQARLSDTANYTCVAKNIVAKRRST 311

Qy        234 SAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSS 293
Db        312 TATVIVYVNGGWSSWAEWSPCSNRCGRGWQKRTRTCTNPAPLNGGAFCEGQ-AFQKTACT 370

Qy        294 LLVSVDGSWSPWSKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCVHS 353
Db        371 TVCPVDGAWTEWSKWSACSTECAHWRSRECMAPPPQNGGRDCSGTLLDSKNCTDGLCVLT 430

Qy        354 ASGPEDVALYVGL-IAVAVCLVLLLLVLILVYCRKKEGLDSDVADSS-ILTSGFQPVSIK 411
Db        431 LETSGDVALYAGLVVAVFVVVAVLMAVGVIVYRRNCRDFDTDITDSSAALTGGFHPVNFK 490

Qy        412 PSKADNPHLL--TIQPDLSTTT-TYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHH 468
Db        491 TARPNNPQLLHPSAPPDLTASAGIYRGPVYALQDS-ADKIPMTNSPLLDPLPSLKIKVYN 549

Qy        469 SS------------------PTSEAEEFVSRLSTQNYFRS----------LPRGTSNMT 499
Db        550 SSTIGSGSGLADGADLLGVLPPGTYPGDF-SRDTHFLHLRSASLGSQHLLGLPRDPSSSV 608

Qy        500 YGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLHKPEDVRLPLA-GCQTLLSPI 558
Db        609 SGTFGCLGGRLSLPGTGVSLLVPNGAIPQGKFYDLYLHINKAEST-LPLSEGSQTVLSPS 667

Qy        559 VSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGSWEQDVLHLGEEAPSHLYYC 618
Db        668 VTCGPTGLLLCRPVVLTVPHCAEVIAGDWIFQLKTQAHQGHWE-EVVTLDEETLNTPCYC 726

Qy        619 QLEASACYVFTEQLGRFALVGEALSVAAAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDA 678
Db        727 QLEAKSCHILLDQLGTYVFMGESYSRSAVKRLQLAIFAPALCTSLEYSLRVYCLEDTPVA 786

Qy        679 LKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWKSKLLVSYQEIPFYHIW 738
Db        787 LKEVLELERTLGGYLVEEPKPLLFKDSYHNLRLSLHDIPHAHWRSKLLAKYQEIPFYHVW 846

Qy        739 NGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESE 798
Db        847 NGSQRALHCTFTLERHSLASTEFTCKVCVRQVEGEGQIFQLHTTLA-ETPAGSLDALCSA 905

Qy        799 AG--VPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLDSHLSFFASKPSP 856
Db        906 PGNAITTQLGPYAFKIPLSIRQKICSSLDAPNSRGNDWRLLAQKLSMDRYLNYFATKASP 965

Qy        857 TAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 899
Db        966 TGVILDLWEARQQDDGDLNSLASALEEMGKSEMLVAMATDGDC 1008

RESULT 10
Q8K1S3
ID   Q8K1S3      PRELIMINARY;      PRT;   945 AA.
AC   Q8K1S3;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Netrin receptor Unc5h2.
GN   UNC5H2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Engelkamp D.;
RT   "Cloning of three mouse unc-5 genes and their expression patterns at
RT   mid-gestation.";
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ487853; CAD32251.1; -.
DR   MGD; MGI:894703; Unc5h2.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0007165; P:signal transduction; IEA.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR000884; TSP1.
DR   InterPro; IPR008085; TSP_1.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00090; tsp_1; 2.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01705; TSP1REPEAT.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00409; IG; 2.
DR   SMART; SM00408; IGc2; 1.
DR   SMART; SM00209; TSP1; 2.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS50092; TSP1; 2.
KW   Immunoglobulin domain; Receptor.
SQ   SEQUENCE   945 AA;  103738 MW;  80E896F0F0E06012 CRC64;

  Query Match             52.2%;  Score 2499;  DB 11;  Length 945;
  Best Local Similarity   52.2%;  Pred. No. 8.4e-222;
  Matches  498;  Conservative  152;  Mismatches  240;  Indels   64;  Gaps   18;

Qy          1 MAVRPGLWPALLGIVLAAW-----LRG--SGAQQSATVANPVPGANPDLLPHFLVEPEDV 53
Db          1 MRARSGVRSALLLALLLCWDPTPSLAGVDSAGQ---VLPDSYPSAPAEQLPYFLLEPQDA 57

Qy         54 YIVKNKPVLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGEPTMEVRINVSRQQV 113
Db         58 YIVKNKPVELHCRAFPATQIYFKCNGEWVSQNDHVTQESLDEATGLRVREVQIEVSRQQV 117

Qy        114 EKVFGLEEYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPP 173
Db        118 EELFGLEDYWCQCVAWSSSGTTKSRRAYIRIAYLRKNFDQEPLAKEVPLDHEVLLQCRPP 177

Qy        174 EGIPPAEVEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSA 233
Db        178 EGVPVAEVEWLKNEDVIDPAQDTNFLLTIDHNLIIRQARLSDTANYTCVAKNIVAKRRST 237

Qy        234 SAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSS 293
Db        238 TATVIVYVNGGWSSWAEWSPCSNRCGRGWQKRTRTCTNPAPLNGGAFCEGQ-AFQKTACT 296

Qy        294 LLVSVDGSWSPWSKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCV-- 351
Db        297 TVCPVDGAWTEWSKWSACSTECAHWRSRECMAPPPQNGGRDCSGTLLDSKNCTDGLCVLN 356

Qy        352 ---------HSASGPEDVALYVGL-IAVAVCLVLLLLVLILVYCRKKEGLDSDVADSS-I 400
Db        357 QRTLNDPKSHPLETSGDVALYAGLVVAVFVVVAVLMAVGVIVYRRNCRDFDTDITDSSAA 416
```

```
Qy    401 LTSGFQPVSIKPSKADNPHLL--TIQPDLSTTT-TYQGSLCPRQDGPSPKFQLTNGHLLS 457
Db    417 LTGGFHPVNFKTARPNNPQLLHPSAPPDLTASAGIYRGPVYALQDS-ADKIPMTNSPLLD 475

Qy    458 PLGGGRHTLHHSS-------------------PTSEAEEFVSRLSTQNYFRS-------- 490
Db    476 PLPSLKIKVYNSSTIGSGSGLADGADLLGVLPPGTYPGDF-SRDTHFLHLRSASLGSQHL 534

Qy    491 --LPRGTSNMTYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLHKPEDVRLPL 548
Db    535 LGLPRDPSSSVSGTFGCLGGRLSLPGTGVSLLVPNGAIPQGKFYDLYLHINKAEST-LPL 593

Qy    549 A-GCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGSWEQDVLHL 607
Db    594 SEGSQTVLSPSVTCGPTGLLLCRPVVLTVPHCAEVIAGDWIFQLKTQAHQGHWE-EVVTL 652

Qy    608 GEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAAKRLKLLLFAPVACTSLEYNI 667
Db    653 DEETLNTPCYCQLEAKSCHILLDQLGTYVFMGESYSRSAVKRLQLAIFAPALCTSLEYSL 712

Qy    668 RVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWKSKLLV 727
Db    713 RVYCLEDTPVALKEVLELERTLGGYLVEEPKPLLFKDSYHNLRLSLHDIPHAHWRSKLLA 772

Qy    728 SYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDT 787
Db    773 KYQEIPFYHVWNGSQRALHCTFTLERHSLASTEFTCKVCVRQVEGEGQIFQLHTTLA-ET 831

Qy    788 RFAELLALESEAG--VPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLDS 845
Db    832 PAGSLDALCSAPGNAITTQLGPYAFKIPLSIRQKICSSLDAPNSRGNDWRLLAQKLSMDR 891

Qy    846 HLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 899
Db    892 YLNYFATKASPTGVILDLWEARQQDDGDLNSLASALEEMGKSEMLVAMATDGDC 945
```

```
RESULT 11
O08722
ID   O08722      PRELIMINARY;      PRT;    945 AA.
AC   O08722;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Transmembrane receptor UNC5H2.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97271897; PubMed=9126742;
RA   Leonardo E.D., Hinck L., Masu M., Keino-Masu K., Ackerman S.L.,
RA   Tessier-Lavigne M.;
RT   "Vertebrate homologues of C. elegans UNC-5 are candidate netrin
RT   receptors.";
RL   Nature 386:833-838(1997).
DR   EMBL; U87306; AAB57679.1; -.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0007165; P:signal transduction; IEA.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR000884; TSP1.
DR   InterPro; IPR008085; TSP_1.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00090; tsp_1; 2.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01705; TSP1REPEAT.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00408; IGc2; 1.
DR   SMART; SM00209; TSP1; 2.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS50092; TSP1; 2.
KW   Immunoglobulin domain; Receptor.
SQ   SEQUENCE   945 AA;  103520 MW;  6E9C2A262E560B9B CRC64;

  Query Match             52.2%;  Score 2497;  DB 11;  Length 945;
  Best Local Similarity   52.0%;  Pred. No. 1.3e-221;
  Matches  501;  Conservative  144;  Mismatches  236;  Indels   82;  Gaps   20;

Qy      1 MAVRPGLWPALLGIVLAAW-----LRG--SGAQQSATVANPVPGANPDLLPHFLVEPEDV 53
Db      1 MRARSGARGALLLALLLCWDPTPSLAGIDSGGQ---ALPDSFPSAPAEQLPHFLLEPEDA 57

Qy     54 YIVKNKPVLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGEPTMEVRINVSRQQV 113
Db     58 YIVKNKPVELHCRAFPATQIYFKCNGEWVSQKGHVTQESLDEATGLRIREVQIEVSRQQV 117

Qy    114 EKVFGLEEYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPP 173
Db    118 EELFGLEDYWCQCVAWSSSGTTKSRRAYIRIAYLRKNFDQEPLAKEVPLDHEVLLQCRPP 177

Qy    174 EGIPPAEVEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSA 233
Db    178 EGVPVAEVEWLKNEDVIDPAQDTNFLLTIDHNLIIRQARLSDTANYTCVAKNIVAKRRST 237

Qy    234 SAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSS 293
Db    238 TATVIVYVNGGWSSWAEWSPCSNRCGRGWQKRTRTCTNPAPLNGGAFCEGQ-ACQKTACT 296

Qy    294 LLVSVDGSWSPWSKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCV-- 351
Db    297 TVCPVDGAWTEWSKWSACSTECAHWRSRECMAPPPQNGGRDCSGTLLDSKNCTDGLCVLN 356

Qy    352 -HSASGPE--------DVALYVGL-IAVAVCLVLLLLVLILVYCRKKEGLDSDVADSS-I 400
Db    357 QRTLNDPKSRPLEPSGDVALYAGLVVAVFVVLAVLMAVGVIVYRRNCRDFDTDITDSSAA 416

Qy    401 LTSGFQPVSIKPSKADNPHLL--TIQPDLSTTT-TYQGSLCPRQDGPSPKFQLTNGHLLS 457
Db    417 LTGGFHPVNFKTARPSNPQLLHPSAPPDLTASAGIYRGPVYALQDS-ADKIPMTNSPLLD 475

Qy    458 PL---------------GGG----------------------RHTLHHSSPTSEAEEFV 479
Db    476 PLPSLKIKVYDSSTIGSGAGLADGADLLGVLPPGTYPGDFSRDTHFLHLRS--------- 526

Qy    480 SRLSTQNYFRSLPRGTSNMTYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLH 539
Db    527 ASLGSQ-HLLGLPRDPSSSVSGTFGCLGGRLTIPGTGVSLLVPNGAIPQGKFYDLYLRIN 585

Qy    540 KPEDVRLPLA-GCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEG 598
Db    586 KTEST-LPLSEGSQTVLSPSVTCGPTGLLLCRPVVLTVPHCAEVIAGDWIFQLKTQAHQG 644

Qy    599 SWEQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAAKRLKLLLFAPV 658
Db    645 HWE-EVVTLDEETLNTPCYCQLEAKSCHILLDQLGTYVFTGESYSRSAVKRLQLAIFAPA 703

Qy    659 ACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPS 718
Db    704 LCTSLEYSLRVYCLEDTPAALKEVLELERTLGGYLVEEPKTLLFKDSYHNLRLSLHDIPH 763

Qy    719 SLWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFS 778
Db    764 AHWRSKLLAKYQEIPFYHVWNGSQKALHCTFTLERHSLASTEFTCKVCVRQVEGEGQIFQ 823

Qy    779 INFNITKDTRFAELLALESEAGVPAL--VGPSAFKIPFLIRQKIISSLDPPCRRGADWRT 836
Db    824 LHTTLA-ETPAGSLDALCSAPGNAATTQLGPYAFKIPLSIRQKICNSLDAPNSRGNDWRL 882
```

```
Qy        837 LAQKLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSE 896
Db        883 LAQKLSMDRYLNYFATKASPTGVILDLWEARQQDDGDLNSLASALEEMGKSEMLVAMTTD 942

Qy        897 AEC 899
Db        943 GDC 945

RESULT 12
Q9D398
ID   Q9D398      PRELIMINARY;      PRT;   945 AA.
AC   Q9D398;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   6330415E02Rik protein.
GN   UNC5H2 OR 6330415E02RIK.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Medulla oblongata;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
DR   EMBL; AK018177; BAB31108.1; -.
DR   MGD; MGI:894703; Unc5h2.
DR   GO; GO:0007165; P:signal transduction; IEA.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR000884; TSP1.
DR   InterPro; IPR008085; TSP_1.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00090; tsp_1; 2.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01705; TSP1REPEAT.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00408; IGc2; 1.
DR   SMART; SM00209; TSP1; 2.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS50092; TSP1; 2.
KW   Immunoglobulin domain.
SQ   SEQUENCE   945 AA;  103725 MW;  43D33B4524E0CBF2 CRC64;

  Query Match             52.1%;  Score 2493;  DB 11;  Length 945;
  Best Local Similarity   52.1%;  Pred. No. 3e-221;
  Matches  497;  Conservative  152;  Mismatches  241;  Indels   64;  Gaps   18;

Qy          1 MAVRPGLWPALLGIVLAAW-----LRG--SGAQQSATVANPVPGANPDLLPHFLVEPEDV 53
Db          1 MRARSGVRSALLLALLLCWDPTPTPSLAGVDSAGQ---VLPDSYPSAPAEQLPYFLLEPQDA 57

Qy         54 YIVKNKPVLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGEPTMEVRINVSRQQV 113
Db         58 YIVKNKPVELHCRAFPATQIYFKCNGEWVSQNDHVTQESLDEATGLRVREVQIEVSRQQV 117

Qy        114 EKVFGLEEYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPP 173
Db        118 EELFGLEDYWCQCVAWSSSGTTKSRRAYIRIAYLRKNFDQEPLAKEVPLDHEVLLQCRPP 177

Qy        174 EGIPPAEVEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSA 233
Db        178 EGVPVAEVEWLKNEDVIDPAQDTNFLLTIDHNLIIRQARLSDTANYTCVAKNIVAKRRST 237

Qy        234 SAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSS 293
Db        238 AATVIVYVNGGWSSWAEWSPCSNRCGRGWQKRTRTCTNPAPLNGGAFCEGQ-AFQKTACT 296

Qy        294 LLVSVDGSWSPWSKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCV-- 351
Db        297 TVCPVDGAWTEWSKWSACSTECAHWRSRECMAPPPQNGGRDCSGTLLDSKNCTDGLCVLN 356

Qy        352 ---------HSASGPEDVALYVGL-IAVAVCLVLLLLVLILVYCRKKEGLDSDVADSS-I 400
Db        357 QRTLNDPKSHPLETSGDVALYAGLVVAVFVVVAVLMAEGVIVYRRNCRDFDTDITDSSAA 416

Qy        401 LTSGFQPVSIKPSKADNPHLL--TIQPDLSTTT-TYQGSLCPRQDGPSPKFQLTNGHLLS 457
Db        417 LTGGFHPVNFKTARPNNPQLLHPSAPPDLTASAGIYRGPVYALQDS-ADKIPMTNSPLLD 475

Qy        458 PLGGGRHTLHHSS-------------------PTSEAEEFVSRLSTQNYFRS-------- 490
Db        476 PLPSLKIKVYNSSTIGSGSGLADGADLLGVLPPGTYPGDF-SRDTHFLHLRSASLGSQHL 534

Qy        491 --LPRGTSNMTYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLHKPEDVRLPL 548
Db        535 LGLPRDPSSSVSGTFGCLGGRLSLPGTGVSLLVPNGAIPQGKFYDLYLHINKAEST-LPL 593

Qy        549 A-GCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGSWEQDVLHL 607
Db        594 SEGSQTVLSPSVTCGPTGLLLCRPVVLTVPHCAEVIAGDWIFQLKTQAHQGHWE-EVVTL 652

Qy        608 GEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAAKRLKLLLFAPVACTSLEYNI 667
Db        653 DEETLNTPCYCQLEAKSCHILLDQLGSYVFMGESYSRSAVKRLQLAIFAPALCTSLEYSL 712

Qy        668 RVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWKSKLLV 727
Db        713 RVYCLEDTPVALKEVLELERTLGGYLVEEPKPLLFKDSYHNLRLSLHDIPHAHWRSKLLA 772

Qy        728 SYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDT 787
Db        773 KYQEIPFYHVWNGSQRALHCTFTLERHSLASTEFTCKVCVRQVEGEGQIFQLHTTLA-ET 831

Qy        788 RFAELLALESEAG--VPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLDS 845
Db        832 PAGSLDALCSAPGNAITTQLGPYAFKIPLSIRQKICSSLDAPDSRGNDWRLLAQKLSMDR 891

Qy        846 HLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 899
Db        892 YLNYFATKASPTGVILDLWEARQQDDGDLNSLASALEEMGKSEMLVAMATDGDC 945

RESULT 13
Q8IZJ1
ID   Q8IZJ1      PRELIMINARY;      PRT;   934 AA.
AC   Q8IZJ1;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
```

```
DE   Transmembrane receptor UNC5H2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22246081; PubMed=12359238;
RA   Komatsuzaki K., Dalvin S., Kinane T.B.;
RT   "Modulation of G(ialpha(2)) signaling by the axonal guidance molecule
RT   UNC5H2.";
RL   Biochem. Biophys. Res. Commun. 297:898-905(2002).
DR   EMBL; AY126437; AAM95701.1; -.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0007165; P:signal transduction; IEA.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR000884; TSP1.
DR   InterPro; IPR008085; TSP_1.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00090; tsp_1; 2.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01705; TSP1REPEAT.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00409; IG; 2.
DR   SMART; SM00408; IGc2; 1.
DR   SMART; SM00209; TSP1; 2.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS50092; TSP1; 2.
KW   Receptor.
SQ   SEQUENCE   934 AA;  102433 MW;  225B3F506D52B780 CRC64;

  Query Match             51.9%;  Score 2486.5;  DB 4;  Length 934;
  Best Local Similarity   52.2%;  Pred. No. 1.2e-220;
  Matches  490;  Conservative  149;  Mismatches  255;  Indels   45;  Gaps   16;

Qy          1 MAVRPGLWPALLGIVLAAW---LRGSGAQQ-SATVANPVPGANPDLLPHFLVEPEDVYIV 56
Db          1 MGARSGARGALLLALLLCWDPRLSQAGTDSGSEVLPDSFPSAPAEPLPYFLQEPQDAYIV 60

Qy         57 KNKPVLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGEPTMEVRINVSRQQVEKV 116
Db         61 KNKPVELRCRAFPATQIYFKCNGEWVSQNDHVTQEGLDEATGLRVREVQIEVSRQQVEEL 120

Qy        117 FGLEEYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGI 176
Db        121 FGLEDYWCQCVAWSSAGTTKSRRAYVRIAYLRKNFDQEPLGKEVPLDHEVLLQCRPPEGV 180

Qy        177 PPAEVEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAA 236
Db        181 PVAEVEWLKNEDVIDPTQDTNFLLTIDHNLIIRQARLSDTANYTCVAKNIVAKRRSTTAT 240

Qy        237 VIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSSLLV 296
Db        241 VIVYVNGGWSSWAEWSPCSNRCGRGWQKRTRTCTNPAPLNGGAFCEGQ-AFQKTACTTIC 299

Qy        297 SVDGSWSPWSKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCVHSASG 356
Db        300 PVDGAWTEWSKWSACSTECAHWRSRECMAPPPQNGGRDCSGTLLDSKNCTDGLCMQMLEA 359

Qy        357 PEDVALYVGL-IAVAVCLVLLLLVLILVYCRKKEGLDSDVADSS-ILTSGFQPVSIKPSK 414
Db        360 SGDAALYAGLVVAIFVVVAILMAVGVVVYRRNCRDFDTDITDSSAALTGGFHPVNFKTAR 419

Qy        415 ADNPHLL--TIQPDLSTTT-TYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSP 471
Db        420 PSNPQLLHPSVPPDLTASAGIYRGPVYALQDS-TDKIPMTNSPLLDPLPSLKVKVYSSST 478

Qy        472 T------SEAEEFVSRLSTQNY------------FRS----------LPRGTSNMTYGTF 503
Db        479 TGSGPGLADGADLLGVLPPGTYPSDFARDTHFLHLRSASLGSQQLLGLPRDPGSSVSGTF 538

Qy        504 NFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLHKPEDVRLPLA-GCQTLLSPIVSCG 562
Db        539 GCLGGRLSIPGTGVSLLVPNGAIPQGKFYEMYLLINKAEST-LPLSEGTQTVLSPSVTCG 597

Qy        563 PPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGSWEQDVLHLGEEAPSHLYYCQLEA 622
Db        598 PTGLLLCRPVILTMPHCAEVSARDWIFQLKTQAHQGHWE-EVVTLDEETLNTPCYCQLEP 656

Qy        623 SACYVFTEQLGRFALVGEALSVAAAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEV 682
Db        657 RACHILLDQLGTYVFTGESYSRSAVKRLQLAVFAPALCTSLEYSLRVYCLEDTPVALKEV 716

Qy        683 VQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWKSKLLVSYQEIPFYHIWNGTQ 742
Db        717 LELERTLGGYLVEEPKPLMFKDSYHNLRLSLHDLPHAHWRSKLLAKYQEIPFYHIWSGSQ 776

Qy        743 RYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAG-- 800
Db        777 KALHCTFTLERHSLASTELTCKICVRQVEGEGQIFQLHTTLA-ETPAGSLDTLCSAPGST 835

Qy        801 VPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLDSHLSFFASKPSPTAMI 860
Db        836 VTTQLGPYAFKIPLSIRQKICNSLDAPNSRGNDWRMLAQKLSMDRYLNYFATKASPTGVI 895

Qy        861 LNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 899
Db        896 LDLWEALQQDDGDLNSLASALEEMGKSEMLVAVATDGDC 934

RESULT 14
Q86SN3
ID   Q86SN3      PRELIMINARY;      PRT;   945 AA.
AC   Q86SN3;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   P53-regulated receptor for death and life.
GN   P53RDL1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22533857; PubMed=12598906;
RA   Tanikawa C., Matsuda K., Fukuda S., Nakamura Y., Arakawa H.;
RT   "p53RDL1 regulates of p53-dependent apoptosis.";
RL   Nat. Cell Biol. 5:216-223(2003).
DR   EMBL; AB096256; BAC57998.1; -.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0007165; P:signal transduction; IEA.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR000884; TSP1.
DR   InterPro; IPR008085; TSP_1.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00090; tsp_1; 2.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01705; TSP1REPEAT.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00409; IG; 2.
DR   SMART; SM00408; IGc2; 1.
DR   SMART; SM00209; TSP1; 2.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
```

```
DR   PROSITE; PS50092; TSP1; 2.
KW   Receptor.
SQ   SEQUENCE   945 AA;  103637 MW;  56064E335F323447 CRC64;

  Query Match             51.8%;  Score 2479;  DB 4;  Length 945;
  Best Local Similarity   51.8%;  Pred. No. 6e-220;
  Matches  493;  Conservative  150;  Mismatches  249;  Indels   60;  Gaps   18;

Qy          1 MAVRPGLWPALLGIVLAAW---LRGSGAQQ-SATVANPVPGANPDLLPHFLVEPEDVYIV 56
Db          1 MGARSGARGALLLALLLCWDPRLSQAGTDSGSEVLPDSFPSAPAEPLPYFLQEPQDAYIV 60

Qy         57 KNKPVLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGEPTMEVRINVSRQQVEKV 116
Db         61 KNKPVELRCRAFPATQIYFKCNGEWVSQNDHVTQEGLDEATGLRVREVQIEVSRQQVEEL 120

Qy        117 FGLEEYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGI 176
Db        121 FGLEDYWCQCVAWSSAGTTKSRRAYVRIAYLRKNFDQEPLGKEVPLDHEVLLQCRPPEGV 180

Qy        177 PPAEVEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAA 236
Db        181 PVAEVEWLKNEDVIDPTQDTNFLLTIDHNLIIRQARLSDTANYTCVAKNIVAKRRSTTAT 240

Qy        237 VIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSSLLV 296
Db        241 VIVYVNGGWSSWAEWSPCSNRCGRGWQKRTRTCTNPAPLNGGAFCEGQ-AFQKTACTTIC 299

Qy        297 SVDGSWSPWSKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCVHS--- 353
Db        300 PVDGAWTEWSKWSACSTECAHWRSRECMAPPPQNGGRDCSGTLLDSKNCTDGLCMQNKKT 359

Qy        354 ----------ASGPEDVALYVGL-IAVAVCLVLLLLVLILVYCRKKEGLDSDVADSS-IL 401
Db        360 LSDPNSHLLEASG--DAALYAGLVVAIFVVVAILMAVGVVVYRRNCRDFDTDITDSSAAL 417

Qy        402 TSGFQPVSIKPSKADNPHLL--TIQPDLSTTT-TYQGSLCPRQDGPSPKFQLTNGHLLSP 458
Db        418 TGGFHPVNFKTARPSNPQLLHPSVPPDLTASAGIYRGPVYALQDS-TDKIPMTNSPLLDP 476

Qy        459 LGGGRHTLHHSSPT------SEAEEFVSRLSTQNY------------FRS---------- 490
Db        477 LPSLKVKVYSSSTTGSGPGLADGADLLGVLPPGTYPSDFARDTHFLHLRSASLGSQQLLG 536

Qy        491 LPRGTSNMTYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLHKPEDVRLPLA- 549
Db        537 LPRDPGSSVSGTFGCLGGRLSIPGTGVSLLVPNGAIPQGKFYEMYLLINKAEST-LPLSE 595

Qy        550 GCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGSWEQDVLHLGE 609
Db        596 GTQTVLSPSVTCGPTGLLLCRPVILTMPHCAEVSARDWIFQLKTQAHQGHWE-EVVTLDE 654

Qy        610 EAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAAKRLKLLLFAPVACTSLEYNIRV 669
Db        655 ETLNTPCYCQLEPRACHILLDQLGTYVFTGESYSRSAVKRLQLAVFAPALCTSLEYSLRV 714

Qy        670 YCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWKSKLLVSY 729
Db        715 YCLEDTPVALKEVLELERTLGGYLVEEPKPLMFKDSYHNLRLSLHDLPHAHWRSKLLAKY 774

Qy        730 QEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRF 789
Db        775 QEIPFYHIWSGSQKALHCTFTLERHSLASTELTCKICVRQVEGEGQIFQLHTTLA-ETPA 833

Qy        790 AELLALESEAG--VPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLDSHL 847
Db        834 GSLDTLCSAPGSTVTTQLGPYAFKIPLSIRQKICNSLDAPNSRGNDWRMLAQKLSMDRYL 893

Qy        848 SFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 899
Db        894 NYFATKASPTGVILDLWEALQQDDGDLNSLASALEEMGKSEMLVAVATDGDC 945
```

```
RESULT 15
Q8K1S2
ID   Q8K1S2     PRELIMINARY;     PRT;   956 AA.
AC   Q8K1S2;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Netrin receptor Unc5h4.
GN   UNC5H4.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Engelkamp D.;
RT   "Cloning of three mouse unc-5 genes and their expression patterns at
RT   mid-gestation.";
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ487854; CAD32252.1; -.
DR   MGD; MGI:2389364; Unc5h4.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0004867; F:serine protease inhibitor activity; IEA.
DR   GO; GO:0007165; P:signal transduction; IEA.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR000215; Serpin.
DR   InterPro; IPR000884; TSP1.
DR   InterPro; IPR008085; TSP_1.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00090; tsp_1; 2.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01705; TSP1REPEAT.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00409; IG; 1.
DR   SMART; SM00408; IGc2; 1.
DR   SMART; SM00209; TSP1; 2.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00284; SERPIN; 1.
DR   PROSITE; PS50092; TSP1; 2.
KW   Immunoglobulin domain; Receptor.
SQ   SEQUENCE   956 AA;  106351 MW;  DFDF07839C10C68D CRC64;

  Query Match             44.8%;  Score 2144.5;  DB 11;  Length 956;
  Best Local Similarity   44.8%;  Pred. No. 5.9e-189;
  Matches  425;  Conservative  162;  Mismatches  283;  Indels   79;  Gaps   16;

Qy          8 WPALLGIVLAAWLRGS----GAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLL 63
Db         15 WLPWLGLFF--WAAGAAAARGADGSEILPDSIPSA-PGTLPHFIEEPEDAYIIKSNPIAL 71

Qy         64 VCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGEPTMEVRINVSRQQVEKVFGLEEYW 123
Db         72 RCKARPAMQIFFKCNGEWVHQNEHVSEESLDESSGLKVREVFINVTRQQVEDFHGPEDYW 131

Qy        124 CQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAEVEW 183
Db        132 CQCVAWSHLGTSKSRKASVRIAYLRKNFEQDPQGREVPIEGMIVLHCRPPEGVPAAEVEW 191

Qy        184 LRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNG 243
Db        192 LKNEEPIDSEQDENIDTRADHNLIIRQARLSDSGNYTCMAANIVAKRRSLSATVVVYVNG 251

Qy        244 GWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDGSWS 303
Db        252 GWSSWTEWSACNVRCGRGWQKRSRTCTNPAPLNGGAFCEGMSVQKITCTA-LCPVDGSWE 310

Qy        304 PWSKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCV----------- 351
```

```
Db    311 VWSEWSVCSPECEHLRIRECTAPPPRNGGKFCEGLSQESENCTDGLCILDKKPLHEIKPQ 370

Qy    352 ----HSASGPEDVALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTSGFQP 407

Db    371 RWSRRGIENASDIALYSGL-GAAVVAVAVLVIGVTLYRRSHSDYGVDVIDSSALTGGFQT 429

Qy    408 VSIKPSKADNPHLL--TIQPDLSTTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLG----- 460

Db    430 FNFKTVRQGNSLLLNPAMQPDLTVSRTYSGPIC-LQD-PLDKELMTESSLFNPLSDIKVK 487

Qy    461 -----------------GGRH------------TLHHSSPTSEAEEFVSRLSTQNYFR 489

Db    488 VQSSFMVSLGVSERAEYHGKNHSGTFPHGNNRGFSTIHPRNKT----PYIQNLS------ 537

Qy    490 SLPRGTSNMTYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLHKPEDVRLPLA 549

Db    538 SLPTRTELRTTGVFGHLGGRLVMPNTGVSLLIPHGAIPEENSWEIYMSINQGEP-SLQSD 596

Qy    550 GCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGSWEQDVLHLGE 609

Db    597 GSEVLLSPEVTCGPPDMLVTTPFALTIPHCADVSSEHWNIHLKKRTQQGKWE-EVMSVED 655

Qy    610 EAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAAKRLKLLLFAPVACTSLEYNIRV 669

Db    656 ESTS--CYCLLDPFACHVLLDSFGTYALTGEPITDCAVKQLKVAVFGCMSCNSLDYNLRV 713

Qy    670 YCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWKSKLLVSY 729

Db    714 YCVDNTPCAFQEVISDERHQGGQLLEEPKLLHFKGNTFSLQVSVLDIPPFLWRIKPFTAC 773

Qy    730 QEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRF 789

Db    774 QEVPFSRVWSSNRQPLHCAFSLERYTPTTTQLSCKICIRQLKGHEQILQVQTSILESERE 833

Qy    790 AELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLDSHLSF 849

Db    834 TITFFAQEDSTFPAQTGPKAFKIPYSIRQRICATFDTPNAKGKDWQMLAQKNSINRNLSY 893

Qy    850 FASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE 898

Db    894 FATQSSPSAVILNLWEARHQQDGDLDSLACALEEIGRTHTKLSNITEPQ 942
```

Search completed: October 6, 2004, 18:19:33
Job time : 89 secs

```
                GenCore version 5.1.6
         Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:          October  6, 2004, 18:20:14 ; Search time 5496 Seconds
                                              (without alignments)
                                              4884.662 Million cell updates/sec

Title:           US-09-970-944-2
Perfect score:   4787
Sequence:        1 MAVRPGLWPALLGIVLAAWL..........AVAGLGQPDAGLFTVSEAEC 899

Scoring table:   BLOSUM62
                 Xgapop 10.0 , Xgapext  0.5
                 Ygapop 10.0 , Ygapext  0.5
                 Fgapop  6.0 , Fgapext  7.0
                 Delop   6.0 , Delext   7.0

Searched:        27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters:      55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09970944/runat_05102004_112010_1636/app_query.fasta_1.1095
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09970944_@CGN_1_1_4087_@runat_05102004_112010_1636 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :       EST:*
                 1:  em_estba:*
                 2:  em_esthum:*
                 3:  em_estin:*
                 4:  em_estmu:*
                 5:  em_estov:*
                 6:  em_estpl:*
                 7:  em_estro:*
                 8:  em_htc:*
                 9:  gb_est1:*
                 10:  gb_est2:*
                 11:  gb_htc:*
                 12:  gb_est3:*
                 13:  gb_est4:*
                 14:  gb_est5:*
                 15:  em_estfun:*
                 16:  em_estom:*
                 17:  em_gss_hum:*
                 18:  em_gss_inv:*
                 19:  em_gss_pln:*
                 20:  em_gss_vrt:*
                 21:  em_gss_fun:*
                 22:  em_gss_mam:*
                 23:  em_gss_mus:*
                 24:  em_gss_pro:*
                 25:  em_gss_rod:*
                 26:  em_gss_phg:*
                 27:  em_gss_vrl:*
                 28:  gb_gss1:*
                 29:  gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 2686 | 56.1 | 3790 | 11 | AK031655 | AK031655 Mus muscu |
| 2 | 2665.5 | 55.7 | 2802 | 29 | AY406491 | AY406491 Homo sapi |
| 3 | 2651.5 | 55.4 | 2791 | 29 | AY406493 | AY406493 Mus muscu |
| 4 | 2493 | 52.1 | 3866 | 11 | AK018177 | AK018177 Mus muscu |
| 5 | 2269 | 47.4 | 2532 | 29 | AY411749 | AY411749 Mus muscu |
| 6 | 2253 | 47.1 | 2532 | 29 | AY411747 | AY411747 Homo sapi |
| 7 | 2169.5 | 45.3 | 2802 | 29 | AY406492 | AY406492 Pan trogl |
| 8 | 2055.5 | 42.9 | 2775 | 29 | AY401471 | AY401471 Mus muscu |
| 9 | 2050.5 | 42.8 | 2775 | 29 | AY401469 | AY401469 Homo sapi |
| 10 | 1890 | 39.5 | 2532 | 29 | AY411748 | AY411748 Pan trogl |
| 11 | 1864.5 | 38.9 | 1532 | 11 | BC033727 | BC033727 Homo sapi |
| 12 | 1550 | 32.4 | 2507 | 29 | AY401470 | AY401470 Pan trogl |
| 13 | 1373.5 | 28.7 | 1034 | 12 | BI758231 | BI758231 603029876 |
| 14 | 1341.5 | 28.0 | 818 | 12 | BI818609 | BI818609 603033362 |
| 15 | 1274 | 26.6 | 874 | 13 | BQ689148 | BQ689148 AGENCOURT |
| 16 | 1268 | 26.5 | 889 | 13 | BQ691915 | BQ691915 AGENCOURT |
| 17 | 1259.5 | 26.3 | 788 | 9 | AI951556 | AI951556 wv36f04.x |
| 18 | 1253 | 26.2 | 788 | 14 | CA317532 | CA317532 UI-M-FW0- |
| c 19 | 1253 | 26.2 | 843 | 13 | BX348193 | BX348193 BX348193 |
| 20 | 1236.5 | 25.8 | 756 | 13 | BU612387 | BU612387 UI-M-EW0- |
| 21 | 1197 | 25.0 | 1175 | 10 | BF530640 | BF530640 602071931 |
| 22 | 1192.5 | 24.9 | 751 | 14 | CF735550 | CF735550 UI-M-HB0- |
| 23 | 1189.5 | 24.8 | 749 | 14 | CF735417 | CF735417 UI-M-HB0- |
| 24 | 1186.5 | 24.8 | 723 | 13 | BU613458 | BU613458 UI-M-EW0- |
| 25 | 1126.5 | 23.5 | 934 | 10 | BF311804 | BF311804 601897316 |
| 26 | 1126 | 23.5 | 909 | 13 | BX364574 | BX364574 BX364574 |
| 27 | 1122 | 23.4 | 678 | 14 | CA749784 | CA749784 UI-M-FD0- |
| 28 | 1112.5 | 23.2 | 859 | 10 | BF311896 | BF311896 601897733 |
| 29 | 1077 | 22.5 | 675 | 14 | CA315487 | CA315487 UI-M-FW0- |
| 30 | 1026 | 21.4 | 1040 | 13 | BX345407 | BX345407 BX345407 |
| 31 | 1024 | 21.4 | 662 | 9 | AL516580 | AL516580 AL516580 |
| 32 | 1022 | 21.3 | 843 | 12 | BG913440 | BG913440 602811321 |
| 33 | 1016.5 | 21.2 | 604 | 10 | BE314370 | BE314370 601147261 |
| 34 | 990.5 | 20.7 | 939 | 13 | BX345406 | BX345406 BX345406 |
| 35 | 989 | 20.7 | 1201 | 13 | BX422753 | BX422753 BX422753 |
| 36 | 934.5 | 19.5 | 720 | 9 | AI959236 | AI959236 fd25g10.y |
| 37 | 924 | 19.3 | 540 | 12 | BI467799 | BI467799 389583 MA |
| 38 | 900 | 18.8 | 824 | 12 | BI737024 | BI737024 603360874 |
| 39 | 879 | 18.4 | 548 | 12 | BI900830 | BI900830 ib81g04.y |
| 40 | 859.5 | 18.0 | 1147 | 13 | BU840446 | BU840446 AGENCOURT |
| 41 | 856.5 | 17.9 | 1111 | 12 | BG298307 | BG298307 602397080 |
| 42 | 849.5 | 17.7 | 987 | 13 | BX452510 | BX452510 BX452510 |
| 43 | 846.5 | 17.7 | 514 | 13 | BX522364 | BX522364 BX522364 |
| 44 | 827 | 17.3 | 815 | 14 | CK140529 | CK140529 AGENCOURT |
| 45 | 814.5 | 17.0 | 608 | 12 | BM487397 | BM487397 pgm2n.pk0 |

ALIGNMENTS

```
RESULT 1
AK031655
LOCUS       AK031655                3790 bp    mRNA    linear   HTC 18-SEP-2003
DEFINITION  Mus musculus 13 days embryo male testis cDNA, RIKEN full-length
            enriched library, clone:6030473H24 product:unc5 homolog (C.
            elegans) 3, full insert sequence.
ACCESSION   AK031655
VERSION     AK031655.1  GI:26327502
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
```

```
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4
  AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the
            FANTOM Consortium.
  TITLE     Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409, 685-690 (2001)
REFERENCE   5
  AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
            Group Phase I & II Team.
  TITLE     Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
REFERENCE   6  (bases 1 to 3790)
  AUTHORS   Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
            Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
            Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
            Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
            Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
            Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
            Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
            Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
            Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
            Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
            Muramatsu,M. and Hayashizaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT     cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
            Please visit our web site for further details.
            URL:http://genome.gsc.riken.go.jp/
            URL:http://fantom.gsc.riken.go.jp/.
FEATURES             Location/Qualifiers
     source          1..3790
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:6030473H24"
                     /db_xref="MGI:2394911"
                     /db_xref="taxon:10090"
                     /clone="6030473H24"
                     /sex="male"
                     /tissue_type="testis"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="13 days embryo"
     CDS             118..2970
                     /note="unnamed protein product; putative
                     unc5 homolog (C. elegans) 3 (MGD|MGI:1095412,
                     GB|NM_009472, evidence: BLASTN, 99%, match=464)"
                     /codon_start=1
                     /protein_id="BAC27495.1"
                     /db_xref="GI:26327503"
                     /translation="MRKGLRATAARCGLGIGYLLQMLVLPALALLSASGTGSAAQDDE
                     FFHELPETFPSDPPEPLPHFLIEPEEAYIVKNKPVNLYCKASPATQIYFKCNSEWVHQ
                     KDHVVDERVDETSGLIVREVSIEISRQQVEELFGPEDYWCQCVAWSSAGTTKSRKAYV
                     RIAYLRKTFEQEPLGKEVSLEQEVLLQCRPPEGIPVAEVEWLKNEDIIDPAEDRNFYI
                     TIDHNLIIKQARLSDTANYTCVAKNIVAKRKSTTATVIVYVNGGWSTWTEWSVCNSRC
                     GRGYQKRTRTCTNPAPLNGGAFCEGQSVQKIACTTLCPVDGRWTSWSKWSTCGTECTH
                     WRRRECTAPAPKNGGKDCDGLVLQSKNCTDGLCMQGFIYPISTEHRPQNEYGFSSAPD
                     SDDVALYVGIVIAVTVCLAITVVVALFVYRKNHRDFESDIIDSSALNGGFQPVNIKAA
                     RQDLLAVPPDLTSAAAMYRGPVYALHDVSDKIPMTNSPILDPLPNLKIKVYNSSGAVT
                     PQDDLAEFSSKLSPQMTQSLLENEALNLKNQSLARQTDPSCTAFGTFNSLGGHLIIPN
                     SGVSLLIPAGAIPQGRVYEMYVTVHRKENMRPPMEDSQTLLTPVVSCGPPGALLTRPV
                     ILTLHHCADPSTEDWKIQLKNQAVQGQWEDVVVVGEENFTTPCYIQLDAEACHILTEN
                     LSTYALVGQSTTKAAAKRLKLAIFGPLCCSSLEYSIRVYCLDDTQDALKEVLQLERQM
                     GGQLLEEPKALRFKGSIHNLRLSIHDIAHSLWKSKLLAKYQEIPFYHIWSGSQRNLHC
                     TFTLERLSLNTVELVCKLCVRQVEGEGQIFQLNCTVSEEPTGIDLPLLDPASTITTVT
                     GPSAFSIPLPIRQKLCSSLDAPQTRGHDWRMLAHKLNLDRYLNYFATKSSPTGVILDL
                     WEAQNFPDGNLSMLAAVLEEMGRHETVVYLAAEGQY"
ORIGIN

Alignment Scores:
Pred. No.:              5.36e-228     Length:         3790
Score:                  2686.00       Matches:        515
Percent Similarity:     71.67%        Conservative:   153
Best Local Similarity:  55.26%        Mismatches:     214
Query Match:            56.11%        Indels:         50
DB:                     11            Gaps:           13

US-09-970-944-2 (1-899) x AK031655 (1-3790)

Qy          9 ProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeuArgGlySerGlyAlaGlnGlnSer 28

Db        193 CCTGCCCTG------GCCCTGCTAAGCGCCAGTGGCACCGGCTCCGCCGCTCAAGATGAT 246

Qy         29 Ala---------ThrValAlaAsnProValProGlyAlaAsnProAspLeuLeuProHis 45

Db        247 GAATTTTTTCACGAACTCCCAGAAACCTTTCCATCTGACCCACCTGAGCCATTGCCACAC 306

Qy         46 PheLeuValGluProGluAspValTyrIleValLysAsnLysProValLeuLeuValCys 65

Db        307 TTCCTCATTGAGCCCGAGGAAGCTTACATTGTGAAGAACAAGCCTGTGAACCTGTATTGT 366

Qy         66 LysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGluTrpValArgGlnVal 85

Db        367 AAAGCCAGCCCTGCCACCCAGATCTACTTCAAGTGCAACAGCGAGTGGGTTCATCAGAAG 426

Qy         86 AspHisValIleGluArgSerThrAspGlySerSerGlyGluProThrMetGluValArg 105

Db        427 GACCACGTAGTAGACGAGAGAGTAGATGAAACCTCTGGTCTAATTGTGAGAGAAGTGAGC 486

Qy        106 IleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGluGluTyrTrpCysGln 125

Db        487 ATTGAGATTTCACGCCAGCAGGTGGAGGAACTGTTTGGGCCTGAAGATTACTGGTGCCAG 546

Qy        126 CysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAlaTyrIleArgIleAla 145

Db        547 TGTGTGGCCTGGAGCTCAGCAGGCACTACGAAGAGTCGGAAGGCATACGTGCGCATTGCG 606

Qy        146 ArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluValSerLeuGluGlnGly 165
```

```
Db    607 TATCTGCGGAAGACATTCGAGCAGGAACCCTTGGGAAAGGAAGTGTCCTTGGAGCAGGAA 666

Qy    166 IleValLeuProCysArgProProGluGlyIleProProAlaGluValGluTrpLeuArg 185
          ::::::|||   |||||||||||||||||||||||||   ||||||||||||||||:::
Db    667 GTCTTACTCCAGTGTCGGCCACCTGAAGGGATCCCAGTGGCTGAGGTAGAATGGCTAAAG 726

Qy    186 AsnGluAspLeuValAspProSerLeuAspProAsnValTyrIleThrArgGluHisSer 205
          ||||||||::::::||||||:::   |||   |||   ||||||||    :::|||:::
Db    727 AATGAAGACATAATTGATCCTGCTGAAGATCGGAACTTTTATATTACTATCGATCACAAC 786

Qy    206 LeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThrCysValAlaLysAsn 225
          |||:::::::::||||||||||||:::|||||||||||||||||||||||||||||||||
Db    787 CTGATCATCAAGCAAGCCCGACTCTCAGATACAGCAAATTATACCTGTGTTGCCAAAAAT 846

Qy    226 IleValAlaArgArgArgSerAlaSerAlaAlaValIleValTyrValAsnGlyGlyTrp 245
          |||||||||:::|||:::|||   :::|||   ||||||||||||||||||||||||||||
Db    847 ATTGTTGCCAAGAGAAAAAGCACCACAGCCACTGTCATCGTGTATGTTAATGGTGGCTGG 906

Qy    246 SerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArgGlyTrpGlnLysArg 265
          |||||||||||||||||||||||||||::::::   ||||||||||||:::|||||||||
Db    907 TCCACCTGGACAGAGTGGTCTGTGTGTAACAGCCGCTGTGGGCGAGGATATCAGAAACGC 966

Qy    266 SerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPheCysGluGlyGlnAsn 285
          :::|||:::||||||||||||||||||||||||||||||||||||||||||||||||:::
Db    967 ACAAGAACCTGCACCAACCCAGCCCCACTCAATGGTGGGGCCTTCTGTGAGGGGCAGAGT 1026

Qy    286 ValHisAspArgThrValSerSerLeuLeuValSerValAspGlySerTrpSerProTrp 305
          |||      :::      :::   |||   |||||||||   |||:::   |||
Db   1027 GTG---CAGAAAATAGCATGCACTACGTTATGTCCAGTGGATGGTAGGTGGACTTCATGG 1083

Qy    306 SerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSerArgGluCysSerAsp 325
          ||||||||||||   ||||||   |||||||||||||||   |||||||||:::
Db   1084 AGCAAATGGTCAACCTGTGGGACTGAATGCACCCACTGGCGCAGGAGGGAGTGTACAGCA 1143

Qy    326 ProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeuAspThrArgAsnCys 345
          |||||||||:::||||||||||::::::|||   |||      |||   ::::::||||||
Db   1144 CCAGCCCCCAAGAACGGGGGTAAGGACTGTGATGGCCTGGTCCTCCAATCCAAGAACTGC 1203

Qy    346 ThrSerAspLeuCysValHis------------------------------------ 352
          |||      ||||||:::
Db   1204 ACTGATGGGCTGTGCATGCAGGGATTCATTTACCCCATTTCAACTGAGCACAGACCCCAG 1263

Qy    353 -----------------SerAlaSerGlyProGluAspValAlaLeuTyrValGlyLeu 366
                           |||||         :::|||||||||||||||||||:::
Db   1264 AATGAATATGGATTTTTCTTCTGCTCCTGACTCAGATGATGTGGCTCTCTACGTGGGGATT 1323

Qy    367 ---IleAlaValAlaValCysLeuValLeuLeuLeuLeuValLeuIleLeuValTyrCys 385
             |||||||||   |||||||||   :::   :::::|||   :::   ||||||
Db   1324 GTGATCGCTGTAACAGTCTGTCTGGCGATCACTGTTGTGGTGGCCCTGTTTGTGTATCGG 1383

Qy    386 ArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIleLeuThrSerGlyPhe 405
          :::         :::||||||:::   |||||||||   |||      ||||||
Db   1384 AAGAACCACCGTGACTTTGAGTCTGACATCATTGACTCCTCAGCACTCAATGGCGGCTTT 1443

Qy    406 GlnProValSerIleLysProSerLysAlaAspAsnProHisLeuLeuThrIleGlnPro 425
          |||||||||:::||||||    ::::::   |||         ||||||   :::   |||
Db   1444 CAGCCTGTGAACATCAAGGCTGCCAGACAAGAT---------CTCCTGGCTGTCCCCCCT 1494

Qy    426 AspLeu---SerThrThrThrThrTyrGlnGlySerLeuCysProArgGlnAspGlyPro 444
          ||||||   |||         |||:::|||   :::               |||
Db   1495 GACCTCACCTCAGCTGCAGCCATGTACAGGGGACCTGTCTATGCTCTGCATGAT---GTC 1551

Qy    445 SerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeuGlyGlyGlyArgHis 464
          |||   |||         :::||||||   :::|||   ||||||         :::
Db   1552 TCAGACAAAATCCCAATGACCAACTCTCCAATTCTGGACCCACTACCCAACTTGAAAATC 1611

Qy    465 ThrLeuHisHisSerSer------------ProThrSerGluAlaGluGluPheValSer 480
             :::::::::||||||            |||      :::      ||||||   |||
Db   1612 AAAGTGTACAACAGCTCAGGTGCTGTCACTCCTCAGGATGACCTTGCCGAGTTCTCATCC 1671

Qy    481 ArgLeuSer---------ThrGlnAsnTyrPhe-------------------------- 488
          :::||||||         ||||||:::
Db   1672 AAACTGTCACCCCAGATGACCCAGTCCTTGCTAGAGAATGAGGCCCTTAACCTGAAGAAC 1731

Qy    489 ArgSerLeuProArgGlyThr------SerAsnMetThrTyrGlyThrPheAsnPheLeu 506
          :::||||||   |||   |||      |||         :::||||||||||||   |||
Db   1732 CAGAGCCTCGCAAGACAGACTGACCCATCCTGCACAGCATTTGGTACCTTCAACTCTCTT 1791

Qy    507 GlyGlyArgLeuMetIleProAsnThrGlyIleSerLeuLeuIleProProAspAlaIle 526
          ||||||   |||:::|||||||||:::|||:::|||||||||||||||      ||||||
Db   1792 GGGGGTCACCTCATCATTCCTAATTCAGGAGTAAGCTTGCTGATTCCCGCTGGGGCCATT 1851

Qy    527 ProArgGlyLysIleTyrGluIleTyrLeuThrLeuHisLysProGluAspValArgLeu 546
          |||:::|||::::::||||||:::|||:::|||:::|||:::      |||::::::|||
Db   1852 CCTCAGGGGAGAGTCTATGAAATGTATGTGACTGTACACAGGAAAGAAAATATGAGGCCC 1911

Qy    547 ProLeuAlaGlyCysGlnThrLeuLeuSerProIleValSerCysGlyProProGlyVal 566
          |||:::            ||||||||||||:::|||:::||||||||||||||||||
Db   1912 CCCATGGAAGACTCTCAGACCCTACTTACCCCTGTGGTGAGCTGTGGGCCTCCTGGAGCT 1971

Qy    567 LeuLeuThrArgProValIleLeuAlaMetAspHisCysGlyGluProSerProAspSer 586
          ||||||||||||||||||||||||      :::   ||||||   :::||||||   :::
Db   1972 CTGCTGACCCGCCCTGTCATCCTCACTCTGCATCACTGTGCAGACCCCAGCACCGAGGAC 2031

Qy    587 TrpSerLeuArgLeuLysLysGlnSerCysGluGlySerTrpGluGlnAspValLeuHis 606
          |||   ::::::||||||   |||:::   :::|||   ||||||   ||||||:::
Db   2032 TGGAAGATCCAGCTCAAAAACCAGGCAGTGCAGGGACAATGGGAG---GATGTTGTGGTG 2088

Qy    607 LeuGlyGluGluAlaProSerHisLeuTyrTyrCysGlnLeuGluAlaSerAlaCysTyr 626
          :::|||||||||      :::         |||   ||||||:::|||   ||||||:::
Db   2089 GTTGGGGAGGAGAACTTCACAACCCCCTGTTACATTCAGCTGGATGCAGAGGCTTGCCAT 2148

Qy    627 ValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGluAlaLeuSerValAlaAla 646
          :::      ||||||   |||      :::||||||||||||::::::   :::   ||||||
Db   2149 ATCCTCACAGAGAACCTCAGTACCTATGCCCTGGTTGGGCAGTCCACCACCAAAGCAGCT 2208

Qy    647 AlaLysArgLeuLysLeuLeuPheAlaProValAlaCysThrSerLeuGluTyrAsn 666
          ||||||||||||||||||      :::|||   |||:::   |||:::|||||||||:::
Db   2209 GCCAAGCGTCTTAAACTGGCCATCTTTGGGCCCCTCTGCTGCTCTTCCCTGGAGTACAGC 2268

Qy    667 IleArgValTyrCysLeuHisAspThrHisAspAlaLeuLysGluValValGlnLeuGlu 686
          ||||||||||||||||||   ||||||   ||||||||||||||||||:::||||||||||
Db   2269 ATTAGAGTCTACTGCCTGGATGACACACAGGATGCCCTGAAGGAAGTTCTACAACTGGAG 2328

Qy    687 LysGlnLeuGlyGlyGlnLeuIleGlnGluProArgValLeuHisPheLysAspSerTyr 706
          :::|||:::||||||||||||::::::||||||:::   |||   ||||||   |||
Db   2329 AGGCAAATGGGAGGACAGCTCCTAGAAGAACCCAAGGCTCTTCGTTTTAAAGGCAGCATC 2388

Qy    707 HisAsnLeuArgLeuSerIleHisAspValProSerSerLeuTrpLysSerLysLeuLeu 726
          ||||||||||||||||||||||||:::      ||||||||||||||||||||||||
Db   2389 CACAACCTGCGCCTGTCTATTCATGACATCGCCCATTCCCTCTGGAAGAGCAAATTGCTG 2448

Qy    727 ValSerTyrGlnGluIleProPheTyrHisIleTrpAsnGlyThrGlnArgTyrLeuHis 746
                |||||||||||||||||||||||||||:::|||:::||||||   ||||||
Db   2449 GCTAAGTATCAGGAAATTCCATTTTACCACATCTGGAGTGGCTCTCAAAGAAACCTCCAC 2508

Qy    747 CysThrPheThrLeuGluArgValSerProSerThrSerAspLeuAlaCysLysLeuTrp 766
          |||||||||||||||||||:::|||   :::|||   :::|||   ||||||||||
Db   2509 TGCACCTTCACTCTGGAAAGACTCAGCCTAAACACAGTGGAACTGGTTTGCAAACTCTGT 2568

Qy    767 ValTrpGlnValGluGlyAspGlyGlnSerPheSerIleAsnPheAsnIleThrLysAsp 786
          |||   ||||||||||||:::||||||   |||   :::|||      ::::::::::::
Db   2569 GTGCGGCAGGTTGAAGGAGAAGGGCAGATCTTCCAGCTCAACTGTACTGTGTCAGAGGAA 2628

Qy    787 ThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAlaGlyValProAlaLeuValGly 806
                   :::|||      |||:::      :::   :::      :::   |||
Db   2629 CCTACTGGCATCGACTTACCTCTCCTGGACCCTGCTAGTACCATCACCACTGTCACCGGA 2688

Qy    807 ProSerAlaPheLysIleProPheLeuIleArgGlnLysIleIleSerSerLeuAspPro 826
          ||||||||||||   ||||||      ||||||||||||:::   ||||||||||||
Db   2689 CCAAGTGCTTTCAGCATTCCTCTCCCTATCCGGCAGAAGCTATGCAGCAGCCTGGATGCC 2748

Qy    827 ProCysArgArgGlyAlaAspTrpArgThrLeuAlaGlnLysLeuHisLeuAspSerHis 846
          |||      ||||||   |||||||||   ||||||   ||||||:::||||||   :::
Db   2749 CCTCAAACAAGAGGCCATGACTGGAGGATGCTGGCCCATAAACTCAACCTGGACAGGTAC 2808
```

```
Qy        847 LeuSerPhePheAlaSerLysProSerProThrAlaMetIleLeuAsnLeuTrpGluAla 866
              |||::::::||||||:::|||   ||||||||    :::||||||:::||||||||||||
Db       2809 TTGAATTACTTTGCCACCAAATCGAGCCCAACTGGCGTAATCCTGGATCTTTGGGAAGCA 2868

Qy        867 ArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAlaAlaValAlaGlyLeuGlyGln 886
              ::::::||||||:::||||||||||||   |||||||||    :::      :::|||:::
Db       2869 CAGAACTTCCCAGATGGAAACCTGAGCATGCTGGCAGCCGTCCTGGAAGAAATGGGAAGA 2928

Qy        887 ProAspAlaGlyLeuPheThrValSerGluAlaGlu 898
                 :::      ::::::      :::|||   :::
Db       2929 CATGAGACAGTGGTGTACTTGGCAGCAGAAGGACAG 2964

RESULT 2
AY406491
LOCUS       AY406491                2802 bp    DNA     linear   GSS 15-DEC-2003
DEFINITION  Homo sapiens UNC5C gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY406491
VERSION     AY406491.1  GI:39762465
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2802)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 2802)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1..2802
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     gene            <1..>2802
                     /gene="UNC5C"
                     /locus_tag="HCM2575"
ORIGIN

Alignment Scores:
Pred. No.:              2.12e-226     Length:          2802
Score:                  2665.50       Matches:         507
Percent Similarity:     72.24%        Conservative:    154
Best Local Similarity:  55.41%        Mismatches:      221
Query Match:            55.68%        Indels:          33
DB:                     29            Gaps:            13

US-09-970-944-2 (1-899) x AY406491 (1-2802)

Qy          9 ProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeuArgGlySerGlyAlaGlnGlnSer 28
              |||||||||             :::|||:::|||          ||||||   ||||||
Db         76 CCTGCCCTG------GCCCTGCTCAGCGCCAGCGGCACTGGCTCCGCCGCCCAAGATGAT 129

Qy         29 ---------AlaThrValAlaAsnProValProGlyAlaAsnProAspLeuLeuProHis 45
                          :::         |||         |||:::   |||||||||
Db        130 GACTTTTTTCATGAACTCCCAGAAACTTTTCCTTCTGATCCACCTGAGCCTCTGCCACAT 189

Qy         46 PheLeuValGluProGluAspValTyrIleValLysAsnLysProValLeuLeuValCys 65
              ||||||:::|||||||||:::   ||||||||||||||||||||||||   |||   |||
Db        190 TTCCTTATTGAGCCTGAAGAAGCTTATATTGTGAAGAATAAGCCTGTGAACCTGTACTGT 249

Qy         66 LysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGluTrpValArgGlnVal 85
              |||||   |||||||||||||||:::|||||||||||||   |||||||||   |||
Db        250 AAAGCAAGCCCTGCCACCCAGATCTATTTCAAGTGTAATAGTGAATGGGTTCATCAGAAG 309

Qy         86 AspHisValIleGluArgSerThrAspGlySerSerGlyGluProThrMetGluValArg 105
              ||||||:::::::::         |||   :::||||||            ||||||
Db        310 GACCACATAGTAGATGAAAGAGTAGATGAAACTTCCGGTCTCATTGTCCGGGAAGTGAGC 369

Qy        106 IleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGluGluTyrTrpCysGln 125
              |||   :::||||||||||||||||||||||:::::::||||||   |||:::||||||||||||
Db        370 ATTGAGATTTCGCGCCAGCAAGTGGAAGAACTCTTTGGACCTGAAGATTACTGGTGCCAG 429

Qy        126 CysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAlaTyrIleArgIleAla 145
              ||||||||||||||||||:::||||||||||||:::||||||||||||:::|||||||||
Db        430 TGTGTGGCCTGGAGCTCCGCGGGTACCACAAAGAGCCGGAAGGCGTATGTGCGCATTGCA 489

Qy        146 ------ArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluValSerLeuGlu 163
                    ||||||   ||||||||||||||||||   ||||||||||||||||||
Db        490 TNNNNNNNNNNNCGGAAGACATTTGAGCAGGAACCCCTAGGAAAGGAAGTGTCTTTGGAA 549

Qy        164 GlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGluValGluTrp 183
              |||   ::::::|||   ||||||||||||||||||||||||   |||||||||||||||
Db        550 CAGGAAGTCTTACTCCAGTGTCGACCACCTGAAGGGATCCCAGTGGCTGAGGTGGAATGG 609

Qy        184 LeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIleThrArgGlu 203
              |||:::|||||||||:::::::||||||      |||   |||   |||||||||   :::
Db        610 TTGAAAAATGAAGACATAATTGATCCCGTTGAAGATCGGAATTTTTATATTACTATTGAT 669

Qy        204 HisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThrCysValAla 223
              |||:::|||::::::::::|||||||||||||:::|||||||||||||||||||||||||
Db        670 CACAACCTCATCATAAAGCAGGCCCGACTCTCTGATACTGCAAATTACACCTGTGTTGCC 729

Qy        224 LysAsnIleValAlaArgArgArgSerAlaSerAlaAlaValIleValTyrValAsnGly 243
              ||||||||||||||:::|||:::|||   :::|||   ||||||||||||||||||||||
Db        730 AAAAACATTGTTGCCAAGAGGAAAAGTACAACTGCCACTGTCATAGTCTATGTCAACGGT 789

Qy        244 GlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArgGlyTrpGln 263
              |||||||||||||||||||||||||||||||||::::::   ||||||||||||:::|||
Db        790 GGCTGGTCCACCTGGACGGAGTGGTCTGTGTGTAACAGCCGCTGTGGACGAGGGTATCAG 849

Qy        264 LysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPheCysGluGly 283
              ||||||:::|||:::|||||||||||||||||||||||||||||||||||||||||||||
Db        850 AAACGTACAAGGACTTGTACCAACCCGGCACCACTCAATGGGGGTGCCTTCTGTGAAGGG 909

Qy        284 GlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGlySerTrpSer 303
              |||:::|||      :::         :::   |||      |||||||||   |||:::
Db        910 CAGAGTGTG---CAGAAAATAGCCTGTACTACGTTATGCCCAGTGGATGGCAGGTGGACG 966

Qy        304 ProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSerArgGluCys 323
              ||||||||||||||||||   ||||||   :::|||||||||||||||   |||||||||
Db        967 CCATGGAGCAAGTGGTCTACTTGTGGAACTGAGTGCACCCACTGGCGCAGGAGGGAGTGC 1026

Qy        324 SerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeuAspThrArg 343
              :::   |||||||||:::||||||||||||::::::|||   |||      |||   ::::::
Db       1027 ACGGCGCCAGCCCCCAAGAATGGAGGCAAGGACTGCGACGGCCTCGTCTTGCAATCCAAG 1086

Qy        344 AsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspValAlaLeuTyr 363
              |||||||||      ||||||:::   :::|||         :::||||||||||||||||
Db       1087 AACTGCACTGATGGGCTTTGCATGCAGACTGCTCCTGATTCAGATGATGTTGCTCTCTAT 1146

Qy        364 ValGlyLeu---IleAlaValAlaValCysLeuValLeuLeuLeuLeuValLeuIleLeu 382
              |||||||||:::   |||||||||   |||||||||   :::   ::::::|||   :::
Db       1147 GTTGGGATTGTGATAGCAGTGATCGTTTGCCTGGCGATCTCTGTAGTTGTGGCCTTGTTT 1206

Qy        383 ValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIleLeuThr 402
              ||||||   :::         :::||||||:::   |||||||||   |||
Db       1207 GTGTATCGGAAGAATCATCGTGACTTTGAGTCAGATATTATTGACTCTTCGGCACTCAAT 1266

Qy        403 SerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeuLeuThr 422
                 |||||||||||||:::||||||   ::::::   |||         ||||||
Db       1267 GGGGGCTTTCAGCCTGTGAACATCAAGGCAGCAAGACAAGAT---------CTGCTGGCT 1317
```

```
Qy    423 IleGlnProAspLeu---SerThrThrThrThrTyrGlnGlySerLeuCysProArgGln 441
Db   1318 GTACCCCCAGACCTCACGTCAGCTGCAGCCATGTACAGAGGACCTGTCTATGCCCTGCAT 1377

Qy    442 AspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeuGlyGly 461
Db   1378 GAC---GTCTCAGACAAAATCCCAATGACCAACTCTCCAATTCTGGATCCACTGCCCAAC 1434

Qy    462 GlyArgHisThrLeuHisHisSerSer------------ProThrSerGluAlaGluGlu 477
Db   1435 CTGAAAATCAAAGTGTACAACACCTCAGGTGCTGTCACCCCCCAAGATGACCTCTCTGAG 1494

Qy    478 PheValSerArgLeuSer---------ThrGlnAsnTyrPhe------------------ 488
Db   1495 TTTACGTCCAAGCTGTCCCCTCAGATGACCCAGTCGTTGTTGGAGAATGAAGCCCTCAGC 1554

Qy    489 ---------ArgSerLeuProArgGlyThr------SerAsnMetThrTyrGlyThrPhe 503
Db   1555 CTGAAGAACCAGAGTCTAGCAAGGCAGACTGATCCATCCTGTACCGCATTTGGCAGCTTC 1614

Qy    504 AsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeuLeuIleProPro 523
Db   1615 AACTCGCTGGGAGGTCACCTTATTGTTCCCAATTCAGGAGTCAGCTTGCTGATTCCCGCT 1674

Qy    524 AspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHisLysProGluAsp 543
Db   1675 GGGGCCATTCCCCAAGGGAGAGTCTACGAAATGTATGTGACTGTACACAGGAAAGAAACT 1734

Qy    544 ValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleValSerCysGlyPro 563
Db   1735 ATGAGGCCACCCATGGATGACTCTCAGACACTTTTGACCCCTGTGGTGAGCTGTGGGCCC 1794

Qy    564 ProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGlyGluProSer 583
Db   1795 CCAGGAGCTCTGCTCACCCGCCCAGTCGTCCTCACTATGCATCACTGCGCAGACCCCAAT 1854

Qy    584 ProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySerTrpGluGlnAsp 603
Db   1855 ACCGAGGACTGGAAAATACTGCTCAAGAACCAGGCAGCACAGGGACAGTGGGAG---GAT 1911

Qy    604 ValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGlnLeuGluAlaSer 623
Db   1912 GTGGTGGTGGTCGGGGAGGAAAACTTCACCACCCCCTGCTACATTCAGCTGGATGCAGAG 1971

Qy    624 AlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGluAlaLeuSer 643
Db   1972 GCCTGCCACATCCTCACAGAGAACCTCAGCACCTACGCCCTGGTAGGACATTCCACCACC 2031

Qy    644 ValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAlaCysThrSerLeu 663
Db   2032 AAAGCGGCTGCGAAGCGCCTCAAGCTGGCCATCTTTGGGCCCCTGTGCTGCTCCTCGCTG 2091

Qy    664 GluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuLysGluValVal 683
Db   2092 GAGTACAGCATCCGAGTCTACTGTCTGGATGACACCCAGGATGCCCTGAAGGAAATTTTA 2151

Qy    684 GlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgValLeuHisPheLys 703
Db   2152 CATCTTGAGAGACAGATGGAGGACAGCTCCTAGAAGAACCTAAGGCTCTTCATTTTAAA 2211

Qy    704 AspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSerLeuTrpLysSer 723
Db   2212 GGCAGCACCCACAACCTGCGCCTGTCAATTCACGATATCGCCCATTCCCTCTGGAAGAGC 2271

Qy    724 LysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsnGlyThrGlnArg 743
Db   2272 AAATTGCTGGCTAAATATCAGGAAATTCCATTTTACCATGTTTGGAGTGGATCTCAAAGA 2331

Qy    744 TyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSerAspLeuAlaCys 763
Db   2332 AACCTGCACTGCACCTTCACTCTGGAAAGATTTAGCCTGAACACAGTGGAGCTGGTTTGC 2391

Qy    764 LysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIleAsnPheAsnIle 783
Db   2392 AAACTCTGTGTGCGGCAGGTGGAAGGAGAAGGGCAGATCTTCCAGCTCAACTGCACCGTG 2451

Qy    784 ThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAlaGlyValProAla 803
Db   2452 TCAGAGGAACCTACTGGCATCGATTTGCCGCTGCTGGATCCTGCGAACACCATCACCACG 2511

Qy    804 LeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLysIleIleSerSer 823
Db   2512 GTCACGGGGCCCAGTGCTTTCAGCATCCCTCTCCCTATCCGGCAGAAGCTCTGTAGCAGC 2571

Qy    824 LeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGlnLysLeuHisLeu 843
Db   2572 CTGGATGCCCCCCAGACGAGAGGCCATGACTGGAGGATGCTGGCCCATAAGCTGAACCTG 2631

Qy    844 AspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMetIleLeuAsnLeu 863
Db   2632 GACAGGTACTTGAATTACTTTGCCACCAAATCCAGCCCAACTGGCGTAATCCTGGATCTT 2691

Qy    864 TrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAlaAlaValAlaGly 883
Db   2692 TGGGAAGCACAGAACTTCCCAGATGGAAACCTGAGCATGCTGGCAGCTGTCTTGGAAGAA 2751

Qy    884 LeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGlu 898
Db   2752 ATGGGAAGACATGAAACGGTGGTGTCCTTAGCAGCAGAAGGGCAG 2796
```

```
RESULT 3
AY406493
LOCUS       AY406493               2791 bp    DNA     linear   GSS 15-DEC-2003
DEFINITION  Mus musculus UNC5C gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY406493
VERSION     AY406493.1  GI:39762467
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 2791)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 2791)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1..2791
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
     gene            <1..>2791
                     /gene="UNC5C"
                     /locus_tag="HCM2575"
ORIGIN

Alignment Scores:
Pred. No.:             3.76e-225      Length:          2791
Score:                 2651.50        Matches:         511
Percent Similarity:    72.49%         Conservative:    153
```

```
Best Local Similarity:  55.79%        Mismatches:   214
Query Match:            55.39%        Indels:        38
DB:                        29         Gaps:          14

US-09-970-944-2 (1-899) x AY406493 (1-2791)

Qy          9 ProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeuArgGlySerGlyAlaGlnGlnSer 28
Db         76 CCTGCCCTG------GCCCTGCTAAGCGCCAGTGGCACCGGCTCCGCCGCTCAAGATGAT 129

Qy         29 Ala---------ThrValAlaAsnProValProGlyAlaAsnProAspLeuLeuProHis 45
Db        130 GAATTTTTTCACGAACTCCCAGAAACCTTTCCATCTGACCCACCTGAGCCATTGCCACAC 189

Qy         46 PheLeuValGluProGluAspValTyrIleValLysAsnLysProValLeuLeuValCys 65
Db        190 TTCCTCATTGAGCCCGAGGAAGCTTACATTGTGAAGAACAAGCCTGTGAACCTGTATTGT 249

Qy         66 LysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGluTrpValArgGlnVal 85
Db        250 AAAGCCAGCCCTGCCACCCAGATCTACTTCAAGTGCAACAGCGAGTGGGTTCATCAGAAG 309

Qy         86 AspHisValIleGluArgSerThrAspGlySerSerGlyGluProThrMetGluValArg 105
Db        310 GACCACGTAGTAGACGAGAGAGTAGATGAAACCTCTGGTCTAATTGTGAGAGAAGTGAGC 369

Qy        106 IleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGluGluTyrTrpCysGln 125
Db        370 ATTGAGATTTCACGCCAGCAGGTGGAGGAACTGTTTGGGCCTGAAGATTACTGGTGCCAG 429

Qy        126 CysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAlaTyrIleArgIleAla 145
Db        430 TGTGTGGCCTGGAGCTCAGCAGGCACTACGAAGAGTCGGAAGGCATACGTGCGCATTGCG 489

Qy        146 ------ArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluValSerLeuGlu 163
Db        490 TTTCCAGATCTGCGGAAGACATTCGAGCAGGAACCCTTGGGAAAGGAAGTGTCCTTGGAG 549

Qy        164 GlnGlyIleValLeuProCysArgProProGlu-GlyIleProProAlaGluValGluTr 183
Db        550 CAGGAAGTCTTACTCCAGTGTCGGCCACCTGAAGGGA------------TCCGTAGAATG 597

Qy        183 pLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIleThrArgGl 203
Db        598 GCTAAAGAATGAAGACATAATTGATCCTGCTGAAGATCGGAACTTTTATATTACTATCGA 657

Qy        203 uHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThrCysValAl 223
Db        658 TCACAACCTGATCATCAAGCAAGCCCGACTCTCAGATACAGCAAATTATACCTGTGTTGC 717

Qy        223 aLysAsnIleValAlaArgArgArgSerAlaSerAlaAlaValIleValTyrValAsnGl 243
Db        718 CAAAAATATTGTTGCCAAGAGAAAAAGCACCACAGCCACTGTCATCGTGTATGTTAATGG 777

Qy        243 yGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArgGlyTrpGl 263
Db        778 TGGCTGGTCCACCTGGACAGAGTGGTCTGTGTGTAACAGCCGCTGTGGGCGAGGATATCA 837

Qy        263 nLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPheCysGluGl 283
Db        838 GAAACGCACAAGAACCTGCACCAACCCAGCCCCACTCAATGGTGGGGCCTTCTGTGAGGG 897

Qy        283 yGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGlySerTrpSe 303
Db        898 GCAGAGTGTG---CAGAAAATAGCATGCACTACGTTATGTCCAGTGGATGGTAGGTGGAC 954

Qy        303 rProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSerArgGluCy 323
Db        955 TTCATGGAGCAAATGGTCAACCTGTGGGACTGAATGCACCCACTGGCGCAGGAGGGAGTG 1014

Qy        323 sSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeuAspThrAr 343
Db       1015 TACAGCACCAGCCCCCAAGAACGGGGGTAAGGACTGTGATGGCCTGGTCCTCCAATCCAA 1074

Qy        343 gAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspValAlaLeuTy 363
Db       1075 GAACTGCACTGATGGGCTGTGCATGCAGGCTGCTCCTGACTCAGATGATGTGGCTCTCTA 1134

Qy        363 rValGlyLeu---IleAlaValAlaValCysLeuValLeuLeuLeuLeuValLeuIleLe 382
Db       1135 CGTGGGGATTGTGATCGCTGTAACAGTCTGTCTGGCGATCACTGTTGTGGTGGCCCTGTT 1194

Qy        382 uValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIleLeuTh 402
Db       1195 TGTGTATCGGAAGAACCACCGTGACTTTGAGTCTGACATCATTGACTCCTCAGCACTCAA 1254

Qy        402 rSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeuLeuTh 422
Db       1255 TGGCGGCTTTCAGCCTGTGAACATCAAGGCTGCCAGACAAGAT---------CTCCTGGC 1305

Qy        422 rIleGlnProAspLeu---SerThrThrThrThrTyrGlnGlySerLeuCysProArgGl 441
Db       1306 TGTCCCCCCTGACCTCACCTCAGCTGCAGCCATGTACAGGGGACCTGTCTATGCTCTGCA 1365

Qy        441 nAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeuGlyGl 461
Db       1366 TGAT---GTCTCAGACAAAATCCCAATGACCAACTCTCCAATTCTGGACCCACTACCCAA 1422

Qy        461 yGlyArgHisThrLeuHisHisSerSer------------ProThrSerGluAlaGluGl 477
Db       1423 CTTGAAAATCAAAGTGTACAACAGCTCAGGTGCTGTCACTCCTCAGGATGACCTTGCCGA 1482

Qy        477 uPheValSerArgLeuSer---------ThrGlnAsnTyrPhe----------------- 488
Db       1483 GTTCTCATCCAAACTGTCACCCCAGATGACCCAGTCCTTGCTAGAGAATGAGGCCCTTAA 1542

Qy        489 ----------ArgSerLeuProArgGlyThr------SerAsnMetThrTyrGlyThrPh 503
Db       1543 CCTGAAGAACCAGAGCCTCGCAAGACAGACTGACCCATCCTGCACAGCATTTGGTACCTT 1602

Qy        503 eAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeuLeuIleProPr 523
Db       1603 CAACTCTCTTGGGGGTCACCTCATCATTCCTAATTCAGGAGTAAGCTTGCTGATTCCCGC 1662

Qy        523 oAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHisLysProGluAs 543
Db       1663 TGGGGCCATTCCTCAGGGGAGAGTCTATGAAATGTATGTGACTGTACACAGGAAAGAAAA 1722

Qy        543 pValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleValSerCysGlyPr 563
Db       1723 TATGAGGCCCCCCATGGAAGACTCTCAGACCCTACTTACCCCTGTGGTGAGCTGTGGGCC 1782

Qy        563 oProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGlyGluProSe 583
Db       1783 TCCTGGAGCTCTGCTGACCCGCCCTGTCATCCTCACTCTGCATCACTGTGCAGACCCCAG 1842

Qy        583 rProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySerTrpGluGlnAs 603
Db       1843 CACCGAGGACTGGAAGATCCAGCTCAAAAACCAGGCAGTGCAGGGACAATGGGAG---GA 1899

Qy        603 pValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGlnLeuGluAlaSe 623
Db       1900 TGTTGTGGTGGTTGGGGAGGAGAACTTCACAACCCCCTGTTACATTCAGCTGGATGCAGA 1959

Qy        623 rAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGluAlaLeuSe 643
Db       1960 GGCTTGCCATATCCTCACAGAGAACCTCAGTACCTATGCCCTGGTTGGGCAGTCCACCAC 2019

Qy        643 rValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAlaCysThrSerLe 663
Db       2020 CAAAGCAGCTGCCAAGCGTCTTAAACTGGCCATCTTTGGGCCCCTCTGCTGCTCTTCCCT 2079

Qy        663 uGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuLysGluValVa 683
Db       2080 GGAGTACAGCATTAGAGTCTACTGCCTGGATGACACACAGGATGCCCTGAAGGAAGTTCT 2139
```

```
Qy        683 lGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgValLeuHisPheLy 703
              :|||||||:::|||:::|||||||||||||::::::||||||:::     ||||||||||
Db       2140 ACAACTGGAGAGGCAAATGGGAGGACAGCTCCTAGAAGAACCCAAGGCTCTTCATTTTAA 2199

Qy        703 sAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSerLeuTrpLysSe 723
              |  |||  ||||||||||||||||||||||||:::        ||||||||||||
Db       2200 AGGCAGCATCCACAACCTGCGCCTGTCTATTCATGACATCGCCCATTCCCTCTGGAAGAG 2259

Qy        723 rLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsnGlyThrGlnAr 743
              |||||||||      ||||||||||||||||||||||||||||:::|||:::|||||
Db       2260 CAAATTGCTGGCTAAGTATCAGGAAATTCCATTTTACCACATCTGGAGTGGCTCTCAAAG 2319

Qy        743 gTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSerAspLeuAlaCy 763
              |  |||||||||||||||||||||||||:::|||    :::|||    :::|||    ||
Db       2320 AAACCTCCACTGCACCTTCACTCTGGAAAGACTCAGCCTAAACACAGTGGAACTGGTTTG 2379

Qy        763 sLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIleAsnPheAsnIl 783
              ||||||   |||   |||||||||||:::||||||  |||   :::|||          ::
Db       2380 CAAACTCTGTGTGCGGCAGGTTGAAGGAGAAGGGCAGATCTTCCAGCTCAACTGTACTGT 2439

Qy        783 eThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAlaGlyValProAl 803
              :::::::::              :::|||      |||:::      :::    :::
Db       2440 GTCAGAGGAACCTACTGGCATCGACTTACCTCTCCTGGACCCTGCTAGTACCATCACCAC 2499

Qy        803 aLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLysIleIleSerSe 823
               :::   ||||||||||||||   ||||||      ||||||||||||:::    ||||
Db       2500 TGTCACCGGACCAAGTGCTTTCAGCATTCCTCTCCCTATCCGGCAGAAGCTATGCAGCAG 2559

Qy        823 rLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGlnLysLeuHisLe 843
              ||||||   |||      ||||||   ||||||||||   ||||||   ||||||:::||
Db       2560 CCTGGATGCCCCTCAAACAAGAGGCCATGACTGGAGGATGCTGGCCCATAAACTCAACCT 2619

Qy        843 uAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMetIleLeuAsnLe 863
              ||||   :::|||:::::::||||||:::|||   |||||||||   :::||||||:::||
Db       2620 GGACAGGTACTTGAATTACTTTGCCACCAAATCGAGCCCAACTGGCGTAATCCTGGATCT 2679

Qy        863 uTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAlaAlaValAlaGl 883
              ||||||||||:::::||||||:::||||||||||||   |||||||||     :::
Db       2680 TTGGGAAGCACAGAACTTCCCAGATGGAAACCTGAGCATGCTGGCAGCCGTCCTGGAAGA 2739

Qy        883 yLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGlu 898
               :::|||:::   :::        :::        :::|||   :::
Db       2740 AATGGGAAGACATGAGACAGTGGTGTCCTTGGCAGCAGAAGGACAG 2785

RESULT 4
AK018177
LOCUS       AK018177                3866 bp    mRNA    linear   HTC 20-SEP-2003
DEFINITION  Mus musculus adult male medulla oblongata cDNA, RIKEN full-length
            enriched library, clone:6330415E02 product:TRANSMEMBRANE RECEPTOR
            UNC5H2 homolog [Rattus norvegicus], full insert sequence.
ACCESSION   AK018177
VERSION     AK018177.1  GI:12857775
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4
  AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the
            FANTOM Consortium.
  TITLE     Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409, 685-690 (2001)
REFERENCE   5
  AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
            Group Phase I & II Team.
  TITLE     Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
REFERENCE   6  (bases 1 to 3866)
  AUTHORS   Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H.,
            Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y.,
            Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K.,
            Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M.,
            Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,
            Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K.,
            Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C.,
            Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
            Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y.,
            Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T.,
            Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K.,
            Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT     Please visit our web site (http://genome.gsc.riken.go.jp/) for
            further details.
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues. First strand cDNA was primed with a primer
            [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was
            prepared by using trehalose thermo-activated reverse transcriptase
            and subsequently enriched for full-length by cap-trapper. cDNA went
            through one round of normalization to Rot = 10.0 and subtraction to
            Rot = 100.0. Second strand cDNA was prepared with the primer
            adapter of sequence [5'
            GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCC 3']. cDNA was cleaved
            with BamHI and XhoI. Vector: a modified pBluescript KS(+) after
            bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3'
            end: BamHI. Host: DH10B.
FEATURES             Location/Qualifiers
     source          1..3866
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:6330415E02"
                     /db_xref="MGI:1910526"
                     /db_xref="taxon:10090"
                     /clone="6330415E02"
                     /sex="male"
                     /tissue_type="medulla oblongata"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="adult"
```

```
CDS             417..3254
                /note="unnamed protein product; TRANSMEMBRANE RECEPTOR
                UNC5H2 homolog [Rattus norvegicus] (SPTR|O08722, evidence:
                FASTY, 96.5%ID, 100%length, match=2835)
                putative"
                /codon_start=1
                /protein_id="BAB31108.1"
                /db_xref="GI:12857776"
                /translation="MRARSGVRSALLLALLLCWDPTPSLAGVDSAGQVLPDSYPSAPA
                EQLPYFLLEPQDAYIVKNKPVELHCRAFPATQIYFKCNGEWVSQNDHVTQESLDEATG
                LRVREVQIEVSRQQVEELFGLEDYWCQCVAWSSSGTTKSRRAYIRIAYLRKNFDQEPL
                AKEVPLDHEVLLQCRPPEGVPVAEVEWLKNEDVIDPAQDTNFLLTIDHNLIIRQARLS
                DTANYTCVAKNIVAKRRSTAATVIVYVNGGWSSWAEWSPCSNRCGRGWQKRTRTCTNP
                APLNGGAFCEGQAFQKTACTTVCPVDGAWTEWSKWSACSTECAHWRSRECMAPPPQNG
                GRDCSGTLLDSKNCTDGLCVLNQRTLNDPKSHPLETSGDVALYAGLVVAVFVVVAVLM
                AEGVIVYRRNCRDFDTDITDSSAALTGGFHPVNFKTARPNNPQLLHPSAPPDLTASAG
                IYRGPVYALQDSADKIPMTNSPLLDPLPSLKIKVYNSSTIGSGSGLADGADLLGVLPP
                GTYPGDFSRDTHFLHLRSASLGSQHLLGLPRDPSSSVSGTFGCLGGRLSLPGTGVSLL
                VPNGAIPQGKFYDLYLHINKAESTLPLSEGSQTVLSPSVTCGPTGLLLCRPVVLTVPH
                CAEVIAGDWIFQLKTQAHQGHWEEVVTLDEETLNTPCYCQLEAKSCHILLDQLGSYVF
                MGESYSRSAVKRLQLAIFAPALCTSLEYSLRVYCLEDTPVALKEVLELERTLGGYLVE
                EPKPLLFKDSYHNLRLSLHDIPHAHWRSKLLAKYQEIPFYHVWNGSQRALHCTFTLER
                HSLASTEFTCKVCVRQVEGEGQIFQLHTTLAETPAGSLDALCSAPGNAITTQLGPYAF
                KIPLSIRQKICSSLDAPDSRGNDWRLLAQKLSMDRYLNYFATKASPTGVILDLWEARQ
                QDDGDLNSLASALEEMGKSEMLVAMATDGDC"
ORIGIN

Alignment Scores:
Pred. No.:              9.95e-211     Length:          3866
Score:                  2493.00       Matches:         497
Percent Similarity:     68.03%        Conservative:    152
Best Local Similarity:  52.10%        Mismatches:      241
Query Match:            52.08%        Indels:          64
DB:                     11            Gaps:            18

US-09-970-944-2 (1-899) x AK018177 (1-3866)

Qy      1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrp--- 19
Db    417 ATGAGGGCCCGGAGCGGGGTGCGGAGCGCGCTGCTGCTGGCGCTGCTGCTTTGCTGGGAT 476

Qy     20 ------------LeuArgGly------SerGlyAlaGlnSerAlaThrValAlaAsn 33
Db    477 CCGACACCGAGCCTAGCAGGCGTTGACTCTGCTGGCCAG---------GTGCTCCCAGAC 527

Qy     34 ProValProGlyAlaAsnProAspLeuLeuProHisPheLeuValGluProGluAspVal 53
Db    528 TCCTACCCATCAGCCCCTGCGGAGCAGCTGCCGTACTTCCTATTGGAGCCACAGGACGCC 587

Qy     54 TyrIleValLysAsnLysProValLeuLeuValCysLysAlaValProAlaThrGlnIle 73
Db    588 TACATCGTAAAGAACAAGCCAGTGGAACTGCACTGCAGAGCCTTCCCTGCCACGCAGATC 647

Qy     74 PhePheLysCysAsnGlyGluTrpValArgGlnValAspHisValIleGluArgSerThr 93
Db    648 TACTTCAAGTGTAATGGCGAGTGGGTCAGCCAGAATGACCACGTCACACAGGAGAGCCTG 707

Qy     94 AspGlySerSerGlyGluProThrMetGluValArgIleAsnValSerArgGlnGlnVal 113
Db    708 GATGAGGCCACAGGCTTGCGGGTGCGAGAGGTGCAGATCGAGGTGTCACGGCAGCAAGTG 767

Qy    114 GluLysValPheGlyLeuGluGluTyrTrpCysGlnCysValAlaTrpSerSerSerGly 133
Db    768 GAGGAACTCTTCGGGCTCGAGGACTACTGGTGCCAGTGCGTGGCCTGGAGCTCTTCGGGA 827

Qy    134 ThrThrLysSerGlnLysAlaTyrIleArgIleAlaArgLeuArgLysAsnPheGluGln 153
Db    828 ACTACCAAGAGTCGCCGAGCCTACATCCGCATTGCCTACTTGCGCAAGAACTTTGACCAG 887

Qy    154 GluProLeuAlaLysGluValSerLeuGluGlnGlyIleValLeuProCysArgProPro 173
Db    888 GAGCCTCTGGCCAAGGAGGTACCCTTGGATCATGAGGTCCTTCTGCAGTGCCGCCCACCG 947

Qy    174 GluGlyIleProProAlaGluValGluTrpLeuArgAsnGluAspLeuValAspProSer 193
Db    948 GAGGGAGTGCCTGTGGCTGAGGTGGAATGGCTCAAGAATGAAGATGTCATTGACCCCGCT 1007

Qy    194 LeuAspProAsnValTyrIleThrArgGluHisSerLeuValValArgGlnAlaArgLeu 213
Db   1008 CAGGACACTAACTTCCTGCTCACCATTGACCACAACCTCATCATCCGCCAGGCGCGCCTC 1067

Qy    214 AlaAspThrAlaAsnTyrThrCysValAlaLysAsnIleValAlaArgArgArgSerAla 233
Db   1068 TCAGACACGGCCAACTACACCTGTGTGGCCAAGAATATCGTGGCCAAGCGCCGGAGCACC 1127

Qy    234 SerAlaAlaValIleValTyrValAsnGlyGlyTrpSerThrTrpThrGluTrpSerVal 253
Db   1128 GCGGCCACAGTCATCGTCTATGTGAATGGAGGCTGGTCCAGCTGGGCAGAGTGGTCACCC 1187

Qy    254 CysSerAlaSerCysGlyArgGlyTrpGlnLysArgSerArgSerCysThrAsnProAla 273
Db   1188 TGTTCCAATCGCTGTGGCCGAGGCTGGCAGAAGCGTACTCGGACCTGCACCAATCCAGCC 1247

Qy    274 ProLeuAsnGlyGlyAlaPheCysGluGlyGlnAsnValHisAspArgThrValSerSer 293
Db   1248 CCACTCAATGGAGGCGCCTTCTGTGAGGGACAG---GCCTTCCAGAAGACAGCTTGCACC 1304

Qy    294 LeuLeuValSerValAspGlySerTrpSerProTrpSerLysTrpSerAlaCysGlyLeu 313
Db   1305 ACCGTGTGCCCAGTGGATGGAGCGTGGACCGAGTGGAGCAAGTGGTCTGCCTGCAGCACA 1364

Qy    314 AspCysThrHisTrpArgSerArgGluCysSerAspProAlaProArgAsnGlyGlyGlu 333
Db   1365 GAGTGTGCGCACTGGCGCAGCCGCGAGTGCATGGCACCGCCACCCCAGAACGGAGGCCGT 1424

Qy    334 GluCysGlnGlyThrAspLeuAspThrArgAsnCysThrSerAspLeuCysVal------ 351
Db   1425 GACTGCAGCGGGACGCTACTTGACTCCAAGAACTGCACTGATGGGCTGTGCGTGCTGAAT 1484

Qy    352 --------------------------HisSerAlaSerGlyProGluAspValAlaLeu 362
Db   1485 CAGAGAACTCTAAACGACCCTAAAAGCCACCCCCTGGAGACATCGGGAGATGTGGCACTG 1544

Qy    363 TyrValGlyLeu---IleAlaValAlaValCysLeuValLeuLeuLeuLeuValLeuIle 381
Db   1545 TACGCAGGCCTTGTGGTGGCCGTCTTTGTGGTGGTAGCGGTTCTCATGGCCGAGGGAGTG 1604

Qy    382 LeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSer---Ile 400
Db   1605 ATCGTATACCGGAGAAACTGCCGGGACTTCGACACGGACATCACCGACTCCTCTGCGGCC 1664

Qy    401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420
Db   1665 CTCACTGGTGGCTTCCACCCTGTCAACTTCAAGACTGCAAGGCCCAACAACCCGCAGCTC 1724

Qy    421 Leu------ThrIleGlnProAspLeuSerThrThrThr---ThrTyrGlnGlySerLeu 437
Db   1725 CTGCACCCGTCCGCCCCTCCAGACCTAACGGCCAGTGCTGGCATCTACCGCGGGCCTGTG 1784

Qy    438 CysProArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSer 457
Db   1785 TATGCCCTGCAGGACTCC---GCCGACAAGATCCCCATGACTAATTCGCCCCTGCTGGAT 1841

Qy    458 ProLeuGlyGlyGlyArgHisThrLeuHisHisSerSer-------------------- 470
Db   1842 CCCCTGCCCAGCCTCAAGATCAAGGTCTATAACTCCAGCACCATCGGTTCTGGGTCTGGC 1901

Qy    471 -----------------------------------ProThrSerGluAlaGluGluPhe 478
Db   1902 CTGGCTGATGGAGCCGACCTGCTGGGTGTCCTCCCGCCGGGCACGTACCCAGGCGATTTC 1961

Qy    479 ValSerArgLeuSerThrGlnAsnTyrPheArgSer----------------------- 490
Db   1962 ---TCCCGGGACACCCATTTCCTGCACCTGCGCAGTGCCAGCCTTGGTTCCCAGCACCTC 2018

Qy    491 ------LeuProArgGlyThrSerAsnMetThrTyrGlyThrPheAsnPheLeuGlyGly 508
```

```
Db      2019 CTGGGCCTACCTCGGGACCCCAGCAGCAGTGTCAGCGGCACCTTTGGTTGCCTGGGAGGA 2078
Qy       509 ArgLeuMetIleProAsnThrGlyIleSerLeuLeuIleProProAspAlaIleProArg 528
Db      2079 AGGCTGAGCCTCCCCGGCACAGGGGTCAGCCTGTTGGTACCAAATGGAGCCATTCCCCAG 2138
Qy       529 GlyLysIleTyrGluIleTyrLeuThrLeuHisLysProGluAspValArgLeuProLeu 548
Db      2139 GGCAAGTTCTATGACCTGTATCTACATATCAACAAGGCCGAAAGCACC---CTCCCACTT 2195
Qy       549 Ala---GlyCysGlnThrLeuLeuSerProIleValSerCysGlyProProGlyValLeu 567
Db      2196 TCAGAAGGTTCCCAGACAGTATTGAGCCCCTCGGTGACCTGTGGGCCCACAGGCCTACTC 2255
Qy       568 LeuThrArgProValIleLeuAlaMetAspHisCysGlyGluProSerProAspSerTrp 587
Db      2256 CTGTGCCGCCCTGTCGTCCTCACCGTGCCCCACTGTGCTGAAGTCATCGCTGGAGACTGG 2315
Qy       588 SerLeuArgLeuLysLysGlnSerCysGluGlySerTrpGluGlnAspValLeuHisLeu 607
Db      2316 ATCTTTCAGCTCAAGACCCAGGCCCATCAGGGCCACTGGGAG---GAGGTGGTGACCTTG 2372
Qy       608 GlyGluGluAlaProSerHisLeuTyrTyrCysGlnLeuGluAlaSerAlaCysTyrVal 627
Db      2373 GATGAGGAGACCCTCAACACACCCTGCTACTGCCAGCTGGAGGCTAAGTCCTGCCACATC 2432
Qy       628 PheThrGluGlnLeuGlyArgPheAlaLeuValGlyGluAlaLeuSerValAlaAlaAla 647
Db      2433 CTGCTGGACCAGCTGGGTTCCTACGTATTCATGGGCGAGTCCTACTCTCGCTCTGCAGTC 2492
Qy       648 LysArgLeuLysLeuLeuLeuPheAlaProValAlaCysThrSerLeuGluTyrAsnIle 667
Db      2493 AAGCGGCTCCAGCTGGCCATCTTCGCCCCAGCCCTCTGCACCTCCCTGGAGTATAGCCTC 2552
Qy       668 ArgValTyrCysLeuHisAspThrHisAspAlaLeuLysGluValValGlnLeuGluLys 687
Db      2553 AGGGTCTACTGTCTGGAGGACACACCTGTAGCACTGAAGGAGGTCCTGGAGCTGGAGAGG 2612
Qy       688 GlnLeuGlyGlyGlnLeuIleGlnGluProArgValLeuHisPheLysAspSerTyrHis 707
Db      2613 ACTCTGGGTGGCTACTTGGTGGAGGAGCCCAAGCCTTTGCTCTTTAAGGACAGTTACCAC 2672
Qy       708 AsnLeuArgLeuSerIleHisAspValProSerSerLeuTrpLysSerLysLeuLeuVal 727
Db      2673 AACCTACGCCTCTCCCTCCATGACATCCCCCATGCCCACTGGAGGAGCAAACTACTGGCC 2732
Qy       728 SerTyrGlnGluIleProPheTyrHisIleTrpAsnGlyThrGlnArgTyrLeuHisCys 747
Db      2733 AAGTACCAGGAGATTCCCTTCTACCACGTCTGGAATGGCAGCCAGAGAGCCCTGCACTGC 2792
Qy       748 ThrPheThrLeuGluArgValSerProSerThrSerAspLeuAlaCysLysLeuTrpVal 767
Db      2793 ACTTTCACCCTGGAGAGGCATAGCCTGGCCTCCACGGAGTTCACCTGTAAGGTCTGCGTG 2852
Qy       768 TrpGlnValGluGlyAspGlyGlnSerPheSerIleAsnPheAsnIleThrLysAspThr 787
Db      2853 CGGCAGGTCGAAGGGGAAGGCCAGATTTTCCAGCTGCACACAACGTTGGCC---GAGACG 2909
Qy       788 ArgPheAlaGluLeuLeuAlaLeuGluSerGluAlaGly------ValProAlaLeuVal 805
Db      2910 CCTGCTGGCTCCCTGGATGCTCTCTGCTCTGCCCCGGGCAATGCCATCACCACCCAGCTG 2969
Qy       806 GlyProSerAlaPheLysIleProPheLeuIleArgGlnLysIleIleSerSerLeuAsp 825
Db      2970 GGACCCTATGCCTTCAAGATACCCCTGTCCATCCGCCAAAAGATCTGCAGCAGCCTGGAC 3029
Qy       826 ProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGlnLysLeuHisLeuAspSer 845
Db      3030 GCCCCCGACTCCCGGGGCAACGACTGGAGGCTGTTGGCGCAGAAGCTGTCCATGGACCGG 3089
Qy       846 HisLeuSerPhePheAlaSerLysProSerProThrAlaMetIleLeuAsnLeuTrpGlu 865
Db      3090 TACCTAAACTACTTCGCCACCAAAGCTAGTCCCACAGGTGTCATCTTAGACCTCTGGGAA 3149
Qy       866 AlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAlaAlaValAlaGlyLeuGly 885
Db      3150 GCTCGGCAACAGGATGACGGGGACCTCAACAGCCTGGCCAGTGCCTTGGAGGAGATGGGC 3209
Qy       886 GlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
Db      3210 AAGAGTGAGATGCTGGTAGCCATGGCCACAGATGGCGATTGC 3251
```

```
RESULT 5
AY411749
LOCUS       AY411749                2532 bp    DNA     linear   GSS 12-DEC-2003
DEFINITION  Mus musculus HCM4327 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY411749
VERSION     AY411749.1  GI:39767717
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 2532)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 2532)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1..2532
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
     gene            <1..>2532
                     /locus_tag="HCM4327"
ORIGIN

Alignment Scores:
Pred. No.:              4.94e-191        Length:         2532
Score:                  2269.00          Matches:        446
Percent Similarity:     69.23%           Conservative:   139
Best Local Similarity:  52.78%           Mismatches:     206
Query Match:            47.40%           Indels:         54
DB:                     29               Gaps:           15

US-09-970-944-2 (1-899) x AY411749 (1-2532)

Qy       103 GluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGluGluTyr 122
Db        13 GAGGTGCAGATCGAGGTGTCACGGCAGCAAGTGGAGGAACTCTTCGGGCTCGAGGACTAC 72
Qy       123 TrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAlaTyrIle 142
Db        73 TGGTGCCAGTGCGTGGCCTGGAGCTCTTCGGGAACTACCAAGAGTCGCCGAGCCTACATC 132
Qy       143 ArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluValSerLeu 162
Db       133 CGCATTGCCTACTTGCGCAAGAACTTTGACCAGGAGCCTCTGGCCAAGGAGGTACCCTTG 192
Qy       163 GluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGluValGlu 182
```

```
Db    193 GATCATGAGGTCCTTCTGCAGTGCCGCCCACCGGAGGGAGTGCCTGTGGCTGAGGTGGAA 252
Qy    183 TrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIleThrArg 202
Db    253 TGGCTCAAGAATGAAGATGTCATTGACCCCGCTCAGGACACTAACTTCCTGCTCACCATT 312
Qy    203 GluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThrCysVal 222
Db    313 GACCACAACCTCATCATCCGCCAGGCGCGCCTCTCAGACACGGCCAACTACACCTGTGTG 372
Qy    223 AlaLysAsnIleValAlaArgArgArgSerAlaSerAlaAlaValIleValTyrValAsn 242
Db    373 GCCAAGAATATCGTGGCCAAGCGCCGGAGCACCACGGCCACAGTCATCGTCTATGTGAAT 432
Qy    243 GlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArgGlyTrp 262
Db    433 GGAGGCTGGTCCAGCTGGGCAGAGTGGTCACCCTGTTCCAATCGCTGTGGCCGAGGCTGG 492
Qy    263 GlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPheCysGlu 282
Db    493 CAGAAGCGTACTCGGACCTGCACCAATCCAGCCCCACTCAATGGAGGCGCCTTCTGTGAG 552
Qy    283 GlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGlySerTrp 302
Db    553 GGACAG---GCCTTCCAGAAGACAGCTTGCACCACCGTGTGCCCAGTGGATGGAGCGTGG 609
Qy    303 SerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSerArgGlu 322
Db    610 ACCGAGTGGAGCAAGTGGTCTGCCTGCAGCACAGAGTGTGCGCACTGGCGCAGCCGCGAG 669
Qy    323 CysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeuAspThr 342
Db    670 TGCATGGCACCGCCACCCCAGAACGGAGGCCGTGACTGCAGCGGGACGCTACTTGACTCC 729
Qy    343 ArgAsnCysThrSerAspLeuCysVal--------------------------------- 351
Db    730 AAGAACTGCACTGATGGGCTGTGCGTGCTGAATCAGAGAACTCTAAACGACCCTAAAAGC 789
Qy    352 HisSerAlaSerGlyProGluAspValAlaLeuTyrValGlyLeu---IleAlaValAla 370
Db    790 CACCCCCTGGAGACATCGGGAGATGTGGCACTGTACGCAGGCCTTGTGGTGGCCGTCTTT 849
Qy    371 ValCysLeuValLeuLeuLeuLeuValLeuIleLeuValTyrCysArgLysLysGluGly 390
Db    850 GTGGTGGTAGCGGTTCTCATGGCCGTGGGAGTGATCGTATACCGGAGAAACTGCCGGGAC 909
Qy    391 LeuAspSerAspValAlaAspSerSer---IleLeuThrSerGlyPheGlnProValSer 409
Db    910 TTCGACACGGACATCACCGACTCCTCTGCGGCCCTCACTGGTGGCTTCCACCCTGTCAAC 969
Qy    410 IleLysProSerLysAlaAspAsnProHisLeuLeu------ThrIleGlnProAspLeu 427
Db    970 TTCAAGACTGCAAGGCCCAACAACCCGCAGCTCCTGCACCCGTCCGCCCCTCCAGACCTA 1029
Qy    428 SerThrThrThr---ThrTyrGlnGlySerLeuCysProArgGlnAspGlyProSerPro 446
Db   1030 ACGGCCAGTGCTGGCATCTACCGCGGGCCTGTGTATGCCCTGCAGGACTCC---GCCGAC 1086
Qy    447 LysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeuGlyGlyGlyArgHisThrLeu 466
Db   1087 AAGATCCCCATGACTAATTCGCCCCTGCTGGATCCCCTGCCCAGCCTCAAGATCAAGGTC 1146
Qy    467 HisHisSerSer------------------------------------------------ 470
Db   1147 TATAACTCCAGCACCATCGGTTCTGGGTCTGGCCTGGCTGATGGAGCCGACCTGCTGGGT 1206
Qy    471 ---------ProThrSerGluAlaGluGluPheValSerArgLeuSerThrGlnAsnTyr 487
Db   1207 GTCCTCCCGCCGGGCACGTACCCAGGCGATTTC---TCCCGGGACACCCATTTCCTGCAC 1263
Qy    488 PheArgSer-----------------------------LeuProArgGlyThrSerAsn 497

Db   1264 CTGCGCAGTGCCAGCCTTGGTTCCCAGCACCTCCTGGGCCTACCTCGGGACCCCAGCAGC 1323
Qy    498 MetThrTyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIle 517
Db   1324 AGTGTCAGCGGCACCTTTGGTTGCCTGGGAGGAAGGCTGAGCCTCCCCGGCACAGGGGTC 1383
Qy    518 SerLeuLeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThr 537
Db   1384 AGCCTGTTGGTACCAAATGGAGCCATTCCCCAGGGCAAGTTCTATGACCTGTATCTACAT 1443
Qy    538 LeuHisLysProGluAspValArgLeuProLeuAla---GlyCysGlnThrLeuLeuSer 556
Db   1444 ATCAACAAGGCCGAAAGCACC---CTCCCACTTTCAGAAGGTTCCCAGACAGTATTGAGC 1500
Qy    557 ProIleValSerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMet 576
Db   1501 CCCTCGGTGACCTGTGGGCCCACAGGCCTACTCCTGTGCCGCCCTGTCGTCCTCACCGTG 1560
Qy    577 AspHisCysGlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCys 596
Db   1561 CCCCACTGTGCTGAAGTCATCGCTGGAGACTGGATCTTTCAGCTCAAGACCCAGGCCCAT 1620
Qy    597 GluGlySerTrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyr 616
Db   1621 CAGGGCCACTGGGAG---GAGGTGGTGACCTTGGATGAGGAGACCCTCAACACACCCTGC 1677
Qy    617 TyrCysGlnLeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAla 636
Db   1678 TACTGCCAGCTGGAGGCTAAGTCCTGCCACATCCTGCTGGACCAGCTGGGTACCTACGTA 1737
Qy    637 LeuValGlyGluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAla 656
Db   1738 TTCATGGGCGAGTCCTACTCTCGCTCTGCAGTCAAGCGGCTCCAGCTGGCCATCTTCGCC 1797
Qy    657 ProValAlaCysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHis 676
Db   1798 CCAGCCCTCTGCACCTCCCTGGAGTATAGCCTCAGGGTCTACTGTCTGGAGGACACACCT 1857
Qy    677 AspAlaLeuLysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGlu 696
Db   1858 GTAGCACTGAAGGAGGTCCTGGAGCTGGAGAGGACTCTGGGTGGCTACTTGGTGGAGGAG 1917
Qy    697 ProArgValLeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspVal 716
Db   1918 CCCAAGCCTTTGCTCTTTAAGGACAGTTACCACAACCTACGCCTCTCCCTCCATGACATC 1977
Qy    717 ProSerSerLeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHis 736
Db   1978 CCCCATGCCCACTGGAGGAGCAAACTACTGGCCAAGTACCAGGAGATTCCCTTCTACCAC 2037
Qy    737 IleTrpAsnGlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerPro 756
Db   2038 GTCTGGAATGGCAGCCAGAGAGCCCTGCACTGCACTTTCACCCTGGAGAGGCATAGCCTG 2097
Qy    757 SerThrSerAspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSer 776
Db   2098 GCCTCCACGGAGTTCACCTGTAAGGTCTGCGTGCGGCAGGTCGAAGGGGAAGGCCAGATT 2157
Qy    777 PheSerIleAsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGlu 796
Db   2158 TTCCAGCTGCACACAACGTTGGCC---GAGACGCCTGCTGGCTCCCTGGATGCTCTCTGC 2214
Qy    797 SerGluAlaGly------ValProAlaLeuValGlyProSerAlaPheLysIleProPhe 814
Db   2215 TCTGCCCCGGGCAATGCCATCACCACCCAGCTGGGACCCTATGCCTTCAAGATACCCCTG 2274
Qy    815 LeuIleArgGlnLysIleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrp 834
Db   2275 TCCATCCGCCAAAAGATCTGCAGCAGCCTGGACGCCCCCAACTCCCGGGGCAACGACTGG 2334
Qy    835 ArgThrLeuAlaGlnLysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysPro 854
Db   2335 AGGCTGTTGGCGCAGAAGCTGTCCATGGACCGGTACCTAAACTACTTCGCCACCAAAGCT 2394
```

```
Qy          855 SerProThrAlaMetIleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeu 874
                ||||||||     :::||||||:::|||||||||||||          :::||||:::|||
Db         2395 AGTCCCACAGGTGTCATCTTAGACCTCTGGGAAGCTCGGCAACAGGATGACGGGGACCTC 2454

Qy          875 SerGlnLeuAlaAlaAlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrVal 894
                :::    ||||||:::|||:::         :::|||:::   :::         :::
Db         2455 AACAGCCTGGCCAGTGCCTTGGAGGAGATGGGCAAGAGTGAGATGCTGGTAGCCATGGCC 2514

Qy          895 SerGluAlaGluCys 899
                ::::::    :::|||
Db         2515 ACAGATGGCGATTGC 2529

RESULT 6
AY411747
LOCUS       AY411747                2532 bp    DNA     linear   GSS 12-DEC-2003
DEFINITION  Homo sapiens HCM4327 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY411747
VERSION     AY411747.1  GI:39767715
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2532)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 2532)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1..2532
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     gene            <1..>2532
                     /locus_tag="HCM4327"
ORIGIN

Alignment Scores:
Pred. No.:              1.33e-189       Length:          2532
Score:                  2253.00         Matches:         443
Percent Similarity:     68.68%          Conservative:    138
Best Local Similarity:  52.36%          Mismatches:      209
Query Match:            47.06%          Indels:          56
DB:                     29              Gaps:            16

US-09-970-944-2 (1-899) x AY411747 (1-2532)

Qy          103 GluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGluGluTyr 122
                ||||||:::|||   ||||||||||||||||||||||::::::|||||||||||||:::|||
Db           13 GAGGTGCAGATCGAGGTGTCGCGGCAGCAGGTGGAGGAGCTCTTTGGGCTGGAGGATTAC 72

Qy          123 TrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAlaTyrIle 142
                ||||||||||||||||||||||||||||:::|||||||||||||||::::::||||||:::
Db           73 TGGTGCCAGTGCGTGGCCTGGAGCTCCGCGGGCACCACCAAGAGTCGCCGAGCCTACGTC 132

Qy          143 ArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluValSerLeu 162
                |||||||||   ||||||||||||||:::|||||||||||      |||||||||   |||
Db          133 CGCATCGCCTACCTGCGCAAGAACTTCGATCAGGAGCCTCTGGGCAAGGAGGTGCCCCTG 192

Qy          163 GluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGluValGlu 182
                :::      ::::::|||   ||||||||||||||||||||:::|||   ||||||||||||
Db          193 GACCATGAGGTTCTCCTGCAGTGCCGCCCGCCGGAGGGGGTGCCTGTGGCCGAGGTGGAA 252

Qy          183 TrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIleThrArg 202
                ||||||:::||||||||||||::::::||||||:::    |||    |||      :::|||
Db          253 TGGCTCAAGAATGAGGATGTCATCGACCCCACCCAGGACACCAACTTCCTGCTCACCATC 312

Qy          203 GluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThrCysVal 222
                :::|||:::|||::::::|||||||||||||||:::||||||||||||||||||||||||
Db          313 GACCACAACCTCATCATCCGCCAGGCCCGCCTGTCGGACACTGCCAACTATACCTGCGTG 372

Qy          223 AlaLysAsnIleValAlaArgArgArgSerAlaSerAlaAlaValIleValTyrValAsn 242
                |||||||||||||||||:::|||||||||      :::|||    |||||||||||||||||
Db          373 GCCAAGAACATCGTGGCCAAACGCCGGAGCACCACTGCCACCGTCATCGTCTACGTGAAT 432

Qy          243 GlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArgGlyTrp 262
                ||||||||||||:::|||      ||||||      |||||||||||||||||||||
Db          433 GGCGGCTGGTCCAGCTGGGCAGAGTGGTCACCCTGCTCCAACCGCTGTGGCCGAGGCTGG 492

Qy          263 GlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPheCysGlu 282
                ||||||||:::|||:::||||||||||||||||||||||||||||||||||||||||||
Db          493 CAGAAGCGCACCCGGACCTGCACCAACCCCGCTCCACTCAACGGAGGGGCCTTCTGCGAG 552

Qy          283 GlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGlySerTrp 302
                ||||||            :::|||       :::    :::       ||||||||:::|||
Db          553 GGCCAG---GCATTCCAGAAGACCGCCTGCACCACCATCTGCCCAGTCGATGGGGCGTGG 609

Qy          303 SerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSerArgGlu 322
                :::   |||||||||||||||||||||         :::|||   ||||||||||||||||
Db          610 ACGGAGTGGAGCAAGTGGTCAGCCTGCAGCACTGAGTGTGCCCACTGGCGTAGCCGCGAG 669

Qy          323 CysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeuAspThr 342
                |||         |||   ||||:::|||||||||   :::|||   ||||||   ||||||:::
Db          670 TGCATGGCGCCCCCACCCAGAACGGAGGCCGTGACTGCAGCGGGACGCTGCTCGACTCT 729

Qy          343 ArgAsnCysThrSerAspLeuCysValHisSer-------------------------- 353
                :::|||||||||      ||||||:::      :::
Db          730 AAGAACTGCACAGATGGGCTGTGCATGCAAAATAAGAAAACTCTAAGCGACCCCAACAGC 789

Qy          354 ------------AlaSerGlyProGluAspValAlaLeuTyrValGlyLeu---IleAla 368
                            |||||||||      |||   |||||||||   ||||||   :::|||
Db          790 CACCTGCTGGAGGCCTCAGGG------GATGCGGCGCTGTATGCGGGGCTCGTGGTGGCC 843

Qy          369 ValAlaValCysLeuValLeuLeuLeuLeuValLeuIleLeuValTyrCysArgLysLys 388
                :::   |||   :::   :::|||:::   |||   ::::::||||||   |||
Db          844 ATCTTCGTGGTCGTGGCAATCCTCATGGCGGTGGGGGTGGTGGTGTACCGCCGCAACTGC 903

Qy          389 GluGlyLeuAspSerAspValAlaAspSerSer---IleLeuThrSerGlyPheGlnPro 407
                      |||:::|||:::   |||||||||      ||||||   ||||||   |||
Db          904 CGTGACTTCGACACAGACATCACTGACTCATCTGCTGCCCTGACTGGTGGTTTCCACCCC 963

Qy          408 ValSerIleLysProSerLysAlaAspAsnProHisLeuLeu------ThrIleGlnPro 425
                |||:::   |||   ::::::      ||||||   ||||||      ::::::   |||
Db          964 GTCAACTTTAAGACGGCAAGGCCCAGCAACCCGCAGCTCCTACACCCCTCTGTGCCTCCT 1023

Qy          426 AspLeuSerThrThrThr---ThrTyrGlnGlySerLeuCysProArgGlnAspGlyPro 444
                ||||||:::   :::      |||:::|||   :::         ||||||
Db         1024 GACCTGACAGCCAGCGCCGGCATCTACCGCGGACCCGTGTATGCCCTGCAGGACTCC--- 1080

Qy          445 SerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeuGlyGlyGlyArgHis 464
                :::   |||   :::||||||      ||||||   ||||||         :::
Db         1081 ACCGACAAAATCCCCATGACCAACTCTCCTCTGCTGGACCCCTTACCCAGCCTTAAGGTC 1140

Qy          465 ThrLeuHisHisSerSerProThr------------------SerGluAlaGluGluPhe 478
                ::::::   ||||||   |||                     ::::::      :::
Db         1141 AAGGTCTACAGCTCCAGCACCACGGGCTCTGGGCCAGGCCTGGCAGATGGGGCTGACCTG 1200

Qy          479 ValSerArgLeuSerThrGlnAsnTyr-------------------------------- 487
                :::      |||         |||
Db         1201 CTGGGGGTCTTGCCGCCTGGCACATACCCTAGCGATTTCGCCCGGGACACCCACTTCCTG 1260
```

```
Qy      488 ---PheArgSer-----------------------------LeuProArgGlyThrSer 496
                  ||||||                             |||||||||
Db     1261 CACCTGCGCAGCGCCAGCCTCGGTTCCCAGCAGCTCTTGGGCCTGCCCCGAGACCCAGGG 1320

Qy      497 AsnMetThrTyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGly 516
            :::          |||||||||        |||||||||||||||   ||||||   ||||||
Db     1321 AGCAGCGTCAGCGGCACCTTTGGCTGCCTGGGTGGGAGGCTCAGCATCCCCGGCACAGGG 1380

Qy      517 IleSerLeuLeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeu 536
            :::|||||||||:::|||     |||||||||:::||||||   ||||||:::||||||
Db     1381 GTCAGCTTGCTGGTGCCCAATGGAGCCATTCCCCAGGGCAAGTTCTACGAGATGTATCTA 1440

Qy      537 ThrLeuHisLysProGluAspValArgLeuProLeuAla---GlyCysGlnThrLeuLeu 555
               ::::::|||   |||         |||||||||:::   |||   ||||||:::|||
Db     1441 CTCATCAACAAGGCAGAAAGTACC---CTCCCGCTTTCAGAAGGGACCCAGACAGTATTG 1497

Qy      556 SerProIleValSerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAla 575
            ||||||   |||:::|||||||||   |||:::||||||   ||||||||||||||||
Db     1498 AGCCCCTCGGTGACCTGTGGACCCACAGGCCTCCTGCTGTGCCGCCCCGTCATCCTCACC 1557

Qy      576 MetAspHisCysGlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSer 595
            |||   ||||||   |||      |||         :::||||||   |||:::
Db     1558 ATGCCCCACTGTGCCGAAGTCAGTGCCCGTGACTGGATCTTTCAGCTCAAGACCCAGGCC 1617

Qy      596 CysGluGlySerTrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeu 615
            :::|||   ||||||   :::|||:::   |||   ||||||        :::
Db     1618 CACCAGGGCCACTGGGAG---GAGGTGGTGACCCTGGATGAGGAGACCCTGAACACACCC 1674

Qy      616 TyrTyrCysGlnLeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPhe 635
               ||||||||||||||       ||||||::::::      :::|||||||||   :::
Db     1675 TGCTACTGCCAGCTGGAGCCCAGGGCCTGTCACATCCTGCTGGACCAGCTGGGCACCTAC 1734

Qy      636 AlaLeuValGlyGluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPhe 655
                     ||||||:::   |||   :::|||   |||||||||:::|||   :::|||
Db     1735 GTGTTCACGGGCGAGTCCTATTCCCGCTCAGCAGTCAAGCGGCTCCAGCTGGCCGTCTTC 1794

Qy      656 AlaProValAlaCysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThr 675
            ||||||         ||||||||||||||||::::::|||||||||||||||   ||||||
Db     1795 GCCCCCGCCCTCTGCACCTCCCTGGAGTACAGCCTCCGGGTCTACTGCCTGGAGGACACG 1854

Qy      676 HisAspAlaLeuLysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGln 695
                  ||||||||||||||::::::||||||:::   |||||||||   |||::::::
Db     1855 CCTGTAGCACTGAAGGAGGTGCTGGAGCTGGAGCGGACTCTGGGCGGATACTTGGTGGAG 1914

Qy      696 GluProArgValLeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAsp 715
            ||||||:::   |||   ||||||||||||||||||||||||||||||:::||||||
Db     1915 GAGCCGAAACCGCTAATGTTCAAGGACAGTTACCACAACCTGCGCCTCTCCCTCCATGAC 1974

Qy      716 ValProSerSerLeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyr 735
            :::|||   :::   |||:::||||||||||||      ||||||||||||||||||||||
Db     1975 CTCCCCCATGCCCATTGGAGGAGCAAGCTGCTGGCCAAATACCAGGAGATCCCCTTCTAT 2034

Qy      736 HisIleTrpAsnGlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSer 755
            |||||||||:::|||:::|||:::   ||||||||||||||||||||||||||   |||
Db     2035 CACATTTGGAGTGGCAGCCAGAAGGCCCTCCACTGCACTTTCACCCTGGAGAGGCACAGC 2094

Qy      756 ProSerThrSerAspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGln 775
               ::::::::::::|||   ||||||:::   |||   ||||||||||||:::||||||
Db     2095 TTGGCCTCCACAGAGCTCACCTGCAAGATCTGCGTGCGGCAAGTGGAAGGGGAGGGCCAG 2154

Qy      776 SerPheSerIleAsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeu 795
               |||   ::::::      :::      :::|||         |||      |||
Db     2155 ATATTCCAGCTGCATACCACTCTGGCA---GAGACACCTGCTGGCTCCCTGGACACTCTC 2211

Qy      796 GluSerGluAlaGly------ValProAlaLeuValGlyProSerAlaPheLysIlePro 813
               |||      |||      |||         :::||||||   ||||||||||||||||
Db     2212 TGCTCTGCCCCTGGCAGCACTGTCACCACCCAGCTGGGACCTTATGCCTTCAAGATCCCA 2271

Qy      814 PheLeuIleArgGlnLysIleIleSerSerLeuAspProProCysArgArgGlyAlaAsp 833
                  ||||||||||||||      :::|||||||||   |||      ||||||   |||
Db     2272 CTGTCCATCCGCCAGAAGATATGCAACAGCCTAGATGCCCCCAACTCACGGGGCAATGAC 2331

Qy      834 TrpArgThrLeuAlaGlnLysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLys 853
            ||||||   ||||||||||||||      :::|||   :::|||::::::||||||:::|||
Db     2332 TGGCGGATGTTAGCACAGAAGCTCTCTATGGACCGGTACCTGAATTACTTTGCCACCAAA 2391

Qy      854 ProSerProThrAlaMetIleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsn 873
                  |||||||||   :::||||||:::||||||||||||          :::|||:::
Db     2392 GCGAGCCCCACGGGTGTGATCCTGGACCTCTGGGAAGCTCTGCAGCAGGACGATGGGGAC 2451

Qy      874 LeuSerGlnLeuAlaAlaAlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThr 893
            |||:::   ||||||:::|||:::      :::|||:::   :::      :::
Db     2452 CTCAACAGCCTGGCGAGTGCCTTGGAGGAGATGGGCAAGAGTGAGATGCTGGTGGCTGTG 2511

Qy      894 ValSerGluAlaGluCys 899
               ::::::   :::|||
Db     2512 GCCACCGACGGGGACTGC 2529
```

RESULT 7
AY406492

```
LOCUS       AY406492                2802 bp    DNA     linear   GSS 15-DEC-2003
DEFINITION  Pan troglodytes UNC5C gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY406492
VERSION     AY406492.1  GI:39762466
KEYWORDS    GSS.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1  (bases 1 to 2802)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 2802)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1..2802
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
     gene            <1..>2802
                     /gene="UNC5C"
                     /locus_tag="HCM2575"
ORIGIN

Alignment Scores:
Pred. No.:              4.65e-182     Length:          2802
Score:                  2169.50       Matches:         438
Percent Similarity:     62.95%        Conservative:    138
Best Local Similarity:  47.87%        Mismatches:      306
Query Match:            45.32%        Indels:          33
DB:                     29            Gaps:            13

US-09-970-944-2 (1-899) x AY406492 (1-2802)

Qy        9 ProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeuArgGlySerGlyAlaGlnGlnSer 28
            |||||||||          :::|||:::|||          ||||||   ||||||
Db       76 CCTGCCCTG------GCCCTGCTCAGCGCCAGCGGCACTGGCTCCGCCGCCCAAGATGAT 129

Qy       29 ---------AlaThrValAlaAsnProValProGlyAlaAsnProAspLeuLeuProHis 45
                     :::               |||            |||:::   |||||||||
```

```
Db    130 GACTTTTTCATGAACTCCCAGAAACTTTTCCTTCGGATCCACCTGAGCCTCTGCCACAT 189
Qy     46 PheLeuValGluProGluAspValTyrIleValLysAsnLysProValLeuLeuValCys 65
Db    190 TTCCTTATTGAACCTGAGGAAGCTTATATTGTGAAGAATAAGCCTGTGAACCTGTACTGT 249
Qy     66 LysAlaValPro------AlaThrGlnIlePhePheLysCysAsnGlyGluTrpValArg 83
Db    250 AAAGCAAGCCCTGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 309
Qy     84 GlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluProThrMetGlu 103
Db    310 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 369
Qy    104 ValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGluGluTyrTrp 123
Db    370 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 429
Qy    124 CysGlnCysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAlaTyrIleArg 143
Db    430 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 489
Qy    144 IleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluValSerLeuGlu 163
Db    490 NNNNNNNNNNNNCGGAAGACATTTGAGCAGGAACCCCTAGGAAAGGAAGTGTCTTTGGAA 549
Qy    164 GlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGluValGluTrp 183
Db    550 CAGGAAGTCTTACTCCAGTGTCGGCCACCTGAAGGGATCCCAGTGGCTGAGNNNNNNNNN 609
Qy    184 LeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIleThrArgGlu 203
Db    610 NNGAAAAATGAAGACATAATTGATCCCGTTGAAGATCGGAATTTTTATATTACTATTGAT 669
Qy    204 HisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThrCysValAla 223
Db    670 CACAACCTCATCATAAAGCAGGCCCGACTCTCTGATACTGCAAATTATACCTGTGTTGCC 729
Qy    224 LysAsnIleValAlaArgArgArgSerAlaSerAlaAlaValIleValTyrValAsnGly 243
Db    730 AAAAACATTGTTGCCAAGAGGAAAAGTACAACTGCCACTGTCATAGTCTATGTCAACGGT 789
Qy    244 GlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArgGlyTrpGln 263
Db    790 GGCTGGTCCACCTGGACGGAGTGGTCTGTGTGTAACAGCCGCTGTGGACGAGGGTATCAG 849
Qy    264 LysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPheCysGluGly 283
Db    850 AAACGTACAAGGACTTGTACCAACCCGGCACCACTCAATGGGGGTGCCTTCTGTGAAGGG 909
Qy    284 GlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGlySerTrpSer 303
Db    910 CAGAGTGTG---CAGAAAATAGCCTGTACTACGTTATGCCCAGTGGATGGCAGGTGGACG 966
Qy    304 ProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSerArgGluCys 323
Db    967 CCATGGAGCAAGTGGTCTACTTGTGGAACTGAGTGCACCCACTGGCGCAGGAGGGAGTGC 1026
Qy    324 SerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeuAspThrArg 343
Db   1027 ACGGCGCCAGCCCCCAAGAATGGAGGCAAGGACTGCGATGGCCTCGTCTTGCAATCCAAG 1086
Qy    344 AsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspValAlaLeuTyr 363
Db   1087 AACTGCACTGATGGGCTTTGCATGCAGACTGCTCCTGATTCAGATGATGTTGCTCTCTAT 1146
Qy    364 ValGlyLeu---IleAlaValAlaValCysLeuValLeuLeuLeuLeuValLeuIleLeu 382
Db   1147 GTTGGGATTGTGATAGCAGTGATCGTTTGCCTGGCGATCTCTGTAGTTGTGGCCTTGTTT 1206
Qy    383 ValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIleLeuThr 402
Db   1207 GTGTATCGGAAGAATCATCGTGACTTTGAGTCAGATATTATTGACTCTTCGGCACTCAAT 1266

Qy    403 SerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeuLeuThr 422
Db   1267 GGGGGCTTTCAGCCTGTGAACATCAAGGCAGCAAGACAAGAT---------CTGCTGGCT 1317
Qy    423 IleGlnProAspLeu---SerThrThrThrThrTyrGlnGlySerLeuCysProArgGln 441
Db   1318 GTACCCCCAGACCTCACGTCAGCTGCAGCCATGTACAGAGGACCTGTCTATGCCCTGCAT 1377
Qy    442 AspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeuGlyGly 461
Db   1378 GAC---GTCTCAGACAAAATCCCAATGACCAACTCTCCAATTCTGGATCCACTGCCCAAC 1434
Qy    462 GlyArgHisThrLeuHisHisSerSer-----------ProThrSerGluAlaGluGlu 477
Db   1435 CTGAAAATCAAAGTGTACAACACCTCAGGTGCTGTCACCCCCCAAGATGACCTCTCTGAG 1494
Qy    478 PheValSerArgLeuSer---------ThrGlnAsnTyrPhe----------------- 488
Db   1495 TTTACGTCCAAGCTGTCCCCTCAGATGACCCAGTCATTGTTGGAGAATGAAGCCCTCAGC 1554
Qy    489 ---------ArgSerLeuProArgGlyThr------SerAsnMetThrTyrGlyThrPhe 503
Db   1555 CTGAAGAACCAGAGTCTAGCAAGGCAGACTGATCCATCCTGTACCGCATTTGGCAGCTTC 1614
Qy    504 AsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeuLeuIleProPro 523
Db   1615 NNNTNGCTGGGNNNNCACCTTATTGTTCCCAATTCAGGAGTCAGCTTGCTGATTCCCGCT 1674
Qy    524 AspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHisLysProGluAsp 543
Db   1675 GGGGCCATTCCCCAAGGGAGAGTCTACGAAATGTATGTGACTGTACACAGGAAAGAAACT 1734
Qy    544 ValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleValSerCysGlyPro 563
Db   1735 ATGAGGCCACCCATGGATGACTCTCAGACACTTTTGACCCCTGTGGTGAGCTGTGGGCCC 1794
Qy    564 ProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGlyGluProSer 583
Db   1795 CCAGGAGCTCTGCTCACCCGCCCCGTCGTCCTTACTATGCATCACTGCGCAGACCCCAAT 1854
Qy    584 ProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySerTrpGluGlnAsp 603
Db   1855 ACCGAGGACTGGAAAATACTGCTCAAGAACCAGGCAGCACAGGGACAGTGGGAG---GAT 1911
Qy    604 ValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGlnLeuGluAlaSer 623
Db   1912 GTGGTGGTGGTCGGGGAGGAAAACTTCACCACCCCCTGCTACATTCAGCTGGATGCAGAG 1971
Qy    624 AlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGluAlaLeuSer 643
Db   1972 GCCTGCCACATCCTCACAGAGAACCTCAGCACCTACGCCCTGGTAGGACATTCCACCACC 2031
Qy    644 ValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAlaCysThrSerLeu 663
Db   2032 AAAGCGGCTGCGAAGCGCCTCAAGCTGGCCATCTTTGGGCCCCTGTGCTGCTCCTCGCTG 2091
Qy    664 GluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuLysGluValVal 683
Db   2092 GAGTACAGCATCCGAGTCTACTGTCTGGATGACACCCAGGATGCCCTGAAGGAAATTTTA 2151
Qy    684 GlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgValLeuHisPheLys 703
Db   2152 CATCTTGAGAGACAGATGGGAGGACAGCTCCTAGAAGAACCTAAGGCTCTTCATTTTAAA 2211
Qy    704 AspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSerLeuTrpLysSer 723
Db   2212 GGCAGCACCCACAACCTGCGCCTGTCAATTCACGATATCGCCCATTCCCTCTGGAAGAGC 2271
Qy    724 LysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsnGlyThrGlnArg 743
Db   2272 AAATTGCTGGCTAAATATCAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2331
```

```
Qy      744 TyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSerAspLeuAlaCys 763
Db     2332 AACCTGCACTGCACCTTCACTCTGGAAAGATTTAGCCTGAACACAGTGGAGCTGGTTTGC 2391

Qy      764 LysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIleAsnPheAsnIle 783
Db     2392 AAACTCTGTGTGCGGCAGGTGGAAGGAGAAGGGCAGATCTTCCAGCTCAACTGCACCGTG 2451

Qy      784 ThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAlaGlyValProAla 803
Db     2452 TCAGAGGAACCTACTGGCATCGATTTGCCGCTGCTGGATCCTGCGAACACCATCACCACG 2511

Qy      804 LeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLysIleIleSerSer 823
Db     2512 GTCACGGGGCCCAGTGCTTTCAGCATCCCTCTCCCTATCCGGCAGAAGCTCTGTAGCAGC 2571

Qy      824 LeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGlnLysLeuHisLeu 843
Db     2572 CTGGATGCCCCCCAGACGAGAGGCCATGACTGGAGGATGCTGGCCCATAAGCNNNNNNNN 2631

Qy      844 AspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMetIleLeuAsnLeu 863
Db     2632 NNNNNGTACTTGAATTACTTTGCCACCAAATCCAGCCCAACTGGCGTAATCCTGGATCTT 2691

Qy      864 TrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAlaAlaValAlaGly 883
Db     2692 TGGGAAGCACAGAACTTCCCAGATGGAAACCTGAGCATGCTGGCAGCTGTCTTGGAAGAA 2751

Qy      884 LeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGlu 898
Db     2752 ATGGGAAGACATGAAACGGTGGTGTCCTTAGCAGCAGAAGGGCAG 2796
```

```
RESULT 8
AY401471
LOCUS       AY401471                2775 bp    DNA     linear   GSS 12-DEC-2003
DEFINITION  Mus musculus HCM0901 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY401471
VERSION     AY401471.1  GI:39757460
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 2775)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 2775)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1..2775
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
     gene            <1..>2775
                     /locus_tag="HCM0901"
ORIGIN

Alignment Scores:
Pred. No.:                  7.1e-172      Length:          2775
Score:                      2055.50       Matches:         407
Percent Similarity:         61.34%        Conservative:    153
Best Local Similarity:      44.58%        Mismatches:      296
Query Match:                42.94%        Indels:          57
DB:                         29            Gaps:            13

US-09-970-944-2 (1-899) x AY401471 (1-2775)
```

```
Qy       24 GlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnProAspLeuLeu 43
Db       49 GGAGCCGATGGCAGTGAAATTCTTCCCGACTCTATCCCATCTGCC---CCTGGAACGCTG 105

Qy       44 ProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysProValLeuLeu 63
Db      106 CCTCATTTCATCGAGGAGCCCGAAGATGCATACATCATCAAGAGCAACCCGATTGCACTC 165

Qy       64 ValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGluTrpValArg 83
Db      166 AGGTGCAAAGCAAGGCCAGCCATGCAGATCTTCTTCAAATGCAATGGGGAGTGGGTCCAT 225

Qy       84 GlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluProThrMetGlu 103
Db      226 CAGAATGAACATGTGTCCGAGGAGAGCCTGGATGAGAGCTCAGGCTTGAAGGTCCGGGAA 285

Qy      104 ValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGluGluTyrTrp 123
Db      286 GTGTTTATCAATGTCACCAGGCAGCAAGTGGAGGACTTCCATGGGCCAGAGGACTACTGG 345

Qy      124 CysGlnCysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAlaTyrIleArg 143
Db      346 TGCCAGTGTGTAGCGTGGAGCCACCTGGGAACGTCCAAGAGTAGGAAGGCATCTGTGCGC 405

Qy      144 IleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluValSerLeuGlu 163
Db      406 ATAGCCTATTTACGGAAAAACTTTGAACAAGATCCACAAGGCAGGGAGGTTCCTATTGAA 465

Qy      164 GlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGluValGluTrp 183
Db      466 GGCATGATCGTGCTGCATTGTCGCCCGCCAGAGGGAGTTCCAGCAGCAGAGGTGGAATGG 525

Qy      184 LeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIleThrArgGlu 203
Db      526 CTGAAAAATGAGGAGCCCATCGACTCCGAACAGGATGAGAACATTGACACCAGGGCTGAC 585

Qy      204 HisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThrCysValAla 223
Db      586 CATAACTTGATCATCAGACAGGCCCGGCTGTCAGACTCAGGGAATTACACCTGTATGGCA 645

Qy      224 LysAsnIleValAlaArgArgArgSerAlaSerAlaAlaValIleValTyrValAsnGly 243
Db      646 GCCAACATCGTGGCCAAGAGGAGGAGCCTCTCAGCAACTGTGGTGGTCTACGTGAATGGA 705

Qy      244 GlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArgGlyTrpGln 263
Db      706 GGCTGGTCTTCCTGGACAGAGTGGTCAGCCTGCAATGTTCGCTGTGGTAGAGGATGGCAG 765

Qy      264 LysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPheCysGluGly 283
Db      766 AAACGTTCCCGGACCTGCACCAACCCCGCTCCTCTCAATGGTGGGGCCTTTTGTGAGGGA 825

Qy      284 GlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGlySerTrpSer 303
Db      826 ATGTCAGTGCAGAAAATAACCTGCACTGCT---CTTTGTCCTGTGGATGGCAGCTGGGAA 882

Qy      304 ProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSerArgGluCys 323
Db      883 GTGTGGAGTGAATGGTCAGTCTGCAGCCCAGAGTGTGAACATCTTCGTATCCGTGAGTGC 942

Qy      324 SerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeuAspThrArg 343
Db      943 ACAGCTCCACCCCCAAGAAATGGGGGCAAGTTCTGTGAAGGTCTCAGCCAGGAATCTGAA 1002
```

```
Qy        344 AsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspValAlaLeuTyr 363
Db       1003 AACTGCACAGATGGTCTCTGCATTCTAGGCATTGAGAATGCCAGCGACATTGCTTTGTAC 1062

Qy        364 ValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuLeuValLeuIleLeuVal 383
Db       1063 TCCGGCTTG---GGTGCAGCCGTGGTAGCTGTTGCGGTCCTGGTCATTGGTGTCACCCTC 1119

Qy        384 TyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIleLeuThrSer 403
Db       1120 TACAGACGGAGCCACAGTGACTACGGAGTGGATGTCATTGACTCTTCTGCATTGACAGGT 1179

Qy        404 GlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeuLeu------ 421
Db       1180 GGCTTCCAGACCTTCAACTTCAAAACAGTTCGTCAAGGAAATTCTCTACTACTGAATCCT 1239

Qy        422 ThrIleGlnProAspLeuSerThrThrThrThrTyrGlnGlySerLeuCysProArgGln 441
Db       1240 GCCATGCAGCCAGATCTGACTGTGAGCCGGACATACAGCGGGCCCATCTGT---CTGCAG 1296

Qy        442 AspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeuGly--- 460
Db       1297 GAT---CCGCTCGACAAGGAGCTCATGACAGAGTCCTCACTCTTCAATCCTTTATCTGAC 1353

Qy        460 ------------------------------------------------------------ 460
Db       1354 ATCAAAGTCAAAGTCCAGAGCTCATTTATGGTTTCCCTGGGAGTGTCTGAGAGAGCGGAG 1413

Qy        461 ------GlyGlyArgHis---------------------------------------Thr 465
Db       1414 TACCATGGCAAGAATCACTCGGGGACTTTCCCCCATGGAAACAACCGCGGATTCAGTACA 1473

Qy        466 LeuHisHisSerSerProThrSerGluAlaGluGluPheValSerArgLeuSerThrGln 485
Db       1474 ATACATCCCAGAAACAAAACG------------CCGTACATCCAAAATCTGTCA------ 1515

Qy        486 AsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThrTyrGlyThrPheAsnPhe 505
Db       1516 ------------TCACTGCCGACAAGGACAGAGCTGAGGACAACTGGTGTCTTTGGCCAT 1563

Qy        506 LeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeuLeuIleProProAspAla 525
Db       1564 TTAGGGGGACGCTTAGTAATGCCAAATACAGNNNNNNNNNNNNNNNNN---NNNNNNNNN 1620

Qy        526 IleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHisLysProGluAspValArg 545
Db       1621 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNC 1680

Qy        546 LeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleValSerCysGlyProProGly 565
Db       1681 CTGCAGTCAGATGGATCTGAGGTTCTCCTGAGTCCTGAAGTCACCTGTGGGCCTCCAGAT 1740

Qy        566 ValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGlyGluProSerProAsp 585
Db       1741 ATGCTTGTCACAACTCCCTTTGCGCTGACCATCCCTCACTGTGCAGACGTCAGTTCAGAG 1800

Qy        586 SerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySerTrpGluGlnAspValLeu 605
Db       1801 CACTGGAACATTCACCTGAAGAAGAGGACCCAGCAGGGGAAGTGGGAG---GAAGTGATG 1857

Qy        606 HisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGlnLeuGluAlaSerAlaCys 625
Db       1858 TCAGTGGAGGATGAGTCCACATCC------TGTTACTGCCTCTTGGATCCCTTTGCCTGT 1911

Qy        626 TyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGluAlaLeuSerValAla 645
Db       1912 CACGTGCTGCTTGATAGCTTTGGAACGTATGCCCTCACAGGAGAGCCAATCACAGACTGT 1971

Qy        646 AlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAlaCysThrSerLeuGluTyr 665
Db       1972 GCCGTAAAGCAGCTCAAGGTGGCTGTGTTTGGCTGCATGTCCTGTAACTCCTTGGATTAC 2031

Qy        666 AsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuLysGluValValGlnLeu 685
Db       2032 AACCTGAGAGTTTACTGTGTGGACAATACTCCTTGTGCATTTCAGGAAGTGATTTCTGAC 2091

Qy        686 GluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgValLeuHisPheLysAspSer 705
Db       2092 GAACGGCACCAAGGCGGCCAGCTCCTGGAAGAACCAAAGCTGCTGCATTTCAAAGGGAAT 2151

Qy        706 TyrHisAsnLeuArgLeuSerIleHisAspValProSerSerLeuTrpLysSerLysLeu 725
Db       2152 ACATTCAGTCTCCAGGTCTCTGTCCTTGACATCCCTCCCTTCCTCTGGAGGATCAAACCG 2211

Qy        726 LeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsnGlyThrGlnArgTyrLeu 745
Db       2212 TTCACTGCATGCCAGGAAGTCCCCTTCTCCCGAGTGTGGAGCAGCAACCGGCAGCCCCTG 2271

Qy        746 HisCysThrPheThrLeuGluArgValSerProSerThrSerAspLeuAlaCysLysLeu 765
Db       2272 CACTGTGCCTTCTCCCTGGAGCGCTACACGCCCACCACCACCCAGCTGTCCTGCAAAATC 2331

Qy        766 TrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIleAsnPheAsnIleThrLys 785
Db       2332 TGCATTCGGCAGCTCAAAGGCCATGAACAGATCCTCCAAGTGCAGACATCCATCCTAGAG 2391

Qy        786 AspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAlaGlyValProAlaLeuVal 805
Db       2392 AGTGAACGAGAAACCATCACTTTCTTCGCACAAGAGGACAGTACCTTTCCTGCACAGACT 2451

Qy        806 GlyProSerAlaPheLysIleProPheLeuIleArgGlnLysIleIleSerSerLeuAsp 825
Db       2452 GGCCCAAAAGCCTTCAAAATTCCCTATTCCATCAGACAGAGGATCTGCGCCACGTTTGAT 2511

Qy        826 ProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGlnLysLeuHisLeuAspSer 845
Db       2512 ACCCCAAATGCCAAAGGCAAGGACTGGCAGATGTTAGCACAGAAGAACAGCATCAATAGG 2571

Qy        846 HisLeuSerPhePheAlaSerLysProSerProThrAlaMetIleLeuAsnLeuTrpGlu 865
Db       2572 AATTTATCATATTTCGCTACCCAAAGCAGCCCTTCCGCTGTCATTTTGAACCTCTGGGAA 2631

Qy        866 AlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAlaAlaValAlaGlyLeuGly 885
Db       2632 GCTCGTCATCAGCAGGATGGGGACCTCGACTCGCTCGCCTGTGCCCTTGAGGAGATTGGG 2691

Qy        886 GlnProAspAlaGlyLeuPheThrValSerGluAlaGlu 898
Db       2692 AGGACACACACGAAACTATCCAACATCACAGAGCCACAG 2730
```

```
RESULT 9
AY401469
LOCUS       AY401469                2775 bp    DNA     linear   GSS 12-DEC-2003
DEFINITION  Homo sapiens HCM0901 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY401469
VERSION     AY401469.1  GI:39757458
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2775)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 2775)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
```

```
TITLE       Direct Submission
JOURNAL     Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1..2775
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     gene            <1..>2775
                     /locus_tag="HCM0901"
ORIGIN

Alignment Scores:
Pred. No.:              1.99e-171      Length:         2775
Score:                  2050.50        Matches:        407
Percent Similarity:     61.28%         Conservative:   158
Best Local Similarity:  44.14%         Mismatches:     306
Query Match:            42.83%         Indels:         52
DB:                     29             Gaps:           12

US-09-970-944-2 (1-899) x AY401469 (1-2775)

Qy     12 LeuGlyIleValLeuAlaAlaTrpLeuArgGlySerGlyAlaGlnGlnSerAlaThrVal 31
Db     14 GTAGGCTGTGTTCTGATGTCTTTTTGAT-TTCACAGGAACTGACAATGGCGAAGCCCTT 72

Qy     32 AlaAsnProValProGlyAlaAsnProAspLeuLeuProHisPheLeuValGluProGlu 51
Db     73 CCCGAATCCATCCCATCAGCT---CCTGGGACACTGCCTCATTTCATAGAGGAGCCAGAT 129

Qy     52 AspValTyrIleValLysAsnLysProValLeuLeuValCysLysAlaValProAlaThr 71
Db    130 GATGCTTATATTATCAAGAGCAACCCTATTGCACTCAGGTGCAAAGCGAGGCCAGCCATG 189

Qy     72 GlnIlePhePheLysCysAsnGlyGluTrpValArgGlnValAspHisValIleGluArg 91
Db    190 CAGATATTCTTCAAATGCAACGGCGAGTGGGTCCATCAGAACGAGCACGTCTCTGAAGAG 249

Qy     92 SerThrAspGlySerSerGlyGluProThrMetGluValArgIleAsnValSerArgGln 111
Db    250 ACTCTGGACGAGAGCTCAGGTTTGAAGGTCCGCGAAGTGTTCATCAATGTTACTAGGCAA 309

Qy    112 GlnValGluLysValPheGlyLeuGluGluTyrTrpCysGlnCysValAlaTrpSerSer 131
Db    310 CAGGTGGAGGACTTCCATGGGCCCGAGGACTATTGGTGCCAGTGTGTGGCGTGGAGCCAC 369

Qy    132 SerGlyThrThrLysSerGlnLysAlaTyrIleArgIleAlaArgLeuArgLysAsnPhe 151
Db    370 CTGGGTACCTCCAAGAGCAGGAAGGCCTCTGTGCGCATAGCCTATTTACGGAAAAACTTT 429

Qy    152 GluGlnGluProLeuAlaLysGluValSerLeuGluGlnGlyIleValLeuProCysArg 171
Db    430 GAACAAGACCCACAAGGAAGGGAAGTTCCCATTGAAGGCATGATTGTACTGCACTGCCGC 489

Qy    172 ProProGluGlyIleProProAlaGluValGluTrpLeuArgAsnGluAspLeuValAsp 191
Db    490 CCACCAGAGGGAGTCCCTGCTGCCGAGGTGGAATGGCTGAAAAATGAAGAGCCCATTGAC 549

Qy    192 ProSerLeuAspProAsnValTyrIleThrArgGluHisSerLeuValValArgGlnAla 211
Db    550 TCTGAACAAGACGAGAACATTGACACCAGGGCTGACCATAACCTGATCATCAGGCAGGCA 609

Qy    212 ArgLeuAlaAspThrAlaAsnTyrThrCysValAlaLysAsnIleValAlaArgArgArg 231
Db    610 CGGCTCTCGGACTCAGGAAATTACACCTGCATGGCAGCCAACATCGTGGCTAAGAGGAGA 669

Qy    232 SerAlaSerAlaAlaValIleValTyrValAsnGlyGlyTrpSerThrTrpThrGluTrp 251
Db    670 AGCCTGTCGGCCACTGTTGTGGTCTACGTGAATGGAGGCTGGTCTTCCTGGACAGAGTGG 729

Qy    252 SerValCysSerAlaSerCysGlyArgGlyTrpGlnLysArgSerArgSerCysThrAsn 271
Db    730 TCAGCCTGCAATGTTCGCTGTGGTAGAGGATGGCAGAAACGTTCCCGGACCTGCACCAAC 789

Qy    272 ProAlaProLeuAsnGlyGlyAlaPheCysGluGlyGlnAsnValHisAspArgThrVal 291
Db    790 CCAGCTCCTCTCAATGGTGGGGCCTTTTGTGAGGGAATGTCAGTGCAGAAAATAACCTGC 849

Qy    292 SerSerLeuLeuValSerValAspGlySerTrpSerProTrpSerLysTrpSerAlaCys 311
Db    850 ACTTCT---CTTTGTCCTGTGGATGGGAGCTGGGAAGTGTGGAGCGAATGGTCCGTCTGC 906

Qy    312 GlyLeuAspCysThrHisTrpArgSerArgGluCysSerAspProAlaProArgAsnGly 331
Db    907 AGTCCAGAGTGTGAACATTTGCGGATCCGGGAGTGCACAGCACCACCCCCGAGAAATGGG 966

Qy    332 GlyGluGluCysGlnGlyThrAspLeuAspThrArgAsnCysThrSerAspLeuCysVal 351
Db    967 GGCAAATTCTGTGAAGGTCTAAGCCAGGAATCTGAAAACTGCACAGATGGTCTTTGCATC 1026

Qy    352 HisSerAlaSerGlyProGluAspValAlaLeuTyrValGlyLeuIleAlaValAlaVal 371
Db   1027 CTAGGCATTGAGAATGCCAGCGACATTGCTTTGTACTCGGGCTTG---GGTGCTGCCGTC 1083

Qy    372 CysLeuValLeuLeuLeuValLeuIleLeuValTyrCysArgLysLysGluGlyLeu 391
Db   1084 GTGGCCGTTGCAGTCCTGGTCATTGGTGTCACCCTTTACAGACGGAGCCAGAGTGACTAT 1143

Qy    392 AspSerAspValAlaAspSerSerIleLeuThrSerGlyPheGlnProValSerIleLys 411
Db   1144 GGCGTGGACGTCATTGACTCTTCTGCATTGACAGGTGGCTTCCAGACCTTCAACTTCAAA 1203

Qy    412 ProSerLysAlaAspAsnProHisLeuLeu------ThrIleGlnProAspLeuSerThr 429
Db   1204 ACAGTCCGTCAAGGTAACTCCCTGCTCCTGAATTCTGCCATGCAGCCAGATCTGACAGTG 1263

Qy    430 ThrThrThrTyrGlnGlySerLeuCysProArgGlnAspGlyProSerProLysPheGln 449
Db   1264 AGCCGGACATACAGCGGACCCATCTGT---CTGCAGGAC---CCTCTGGACAAGGAGCTC 1317

Qy    450 LeuThrAsnGlyHisLeuLeuSerProLeuGly------------------------- 460
Db   1318 ATGACAGAGTCCTCACTCTTTAACCCTTTGTCGGACATCAAAGTGAAAGTCCAGAGCTCG 1377

Qy    460 ------------------------------------------------------------ 460
Db   1378 TTCATGGTTTCCCTGGGAGTGTCTGAGAGAGCTGAGTACCACGGCAAGAATCATTCCAGG 1437

Qy    461 ------------GlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGlu 476
Db   1438 ACTTTTCCCCATGGAAACAACCACAGCTTTAGTACAATGCATCCCAGAAATAAAATG--- 1494

Qy    477 GluPheValSerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSer 496
Db   1495 CCCTACATCCAAAATCTGTCA------------------TCACTCCCCACAAGGACAGAA 1536

Qy    497 AsnMetThrTyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGly 516
Db   1537 CTGAGGACAACTGGTGTCTTTGGCCATTTAGGGGGGCGCTTAGTAATGCCAAATACAGNN 1596

Qy    517 IleSerLeuLeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeu 536
Db   1597 NNNNNNNNNNNNNNN---NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1653

Qy    537 ThrLeuHisLysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSer 556
Db   1654 NNNNNNNNNNNNNNNNNNNNNNNNNNNCCTCCAGTCAGATGGCTCTGAGGTGCTCCTGAGT 1713

Qy    557 ProIleValSerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMet 576
Db   1714 CCTGAAGTCACCTGTGGTCCTCCAGACATGATCGTCACCACTCCCTTTGCATTGACCATC 1773

Qy    577 AspHisCysGlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCys 596
```

```
Db       1774 CCGCACTGTGCAGATGTCAGTTCTGAGCATTGGAATATCCATTTAAAGAAGAGGACACAG 1833

Qy        597 GluGlySerTrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyr 616
              :::|||  |||||  :::|||:::  :::  :::|||:::  |||
Db       1834 CAGGGCAAATGGGAG---GAAGTGATGTCAGTGGAAGATGAATCTACATCC------TGT 1884

Qy        617 TyrCysGlnLeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAla 636
              |||||  |||:::  ||||||:::|||  :::  |||  :::|||
Db       1885 TACTGCCTTTTGGACCCCTTTGCGTGTCATGTGCTCCTGGACAGCTTTGGGACCTATGCG 1944

Qy        637 LeuValGlyGluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAla 656
              |||  ||||||  ::::::  |||  |||:::||||||:::  :::|||
Db       1945 CTCACTGGAGAGCCAATCACAGACTGTGCCGTGAAGCAACTGAAGGTGGCGGTTTTTGGC 2004

Qy        657 ProValAlaCysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHis 676
              ::::::|||  |||||:::||||||:::||||||||||||:::  :::|||
Db       2005 TGCATGTCCTGTAACTCCCTGGATTACAACTTGAGAGTTTACTGTGTGGACAATACCCCT 2064

Qy        677 AspAlaLeuLysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGlu 696
              |||  :::|||||||||  |||:::  ||||||||||||::::::|||
Db       2065 TGTGCATTTCAGGAAGTGGTTTCAGATGAAAGGCATCAAGGTGGACAGCTCCTGGAAGAA 2124

Qy        697 ProArgValLeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspVal 716
              |||::::::||||||||||||  :::  :::|||::::::|||:::  |||:::
Db       2125 CCAAAATTGCTGCATTTCAAAGGGAATACCTTTAGTCTTCAGATTTCTGTCCTTGATATT 2184

Qy        717 ProSerSerLeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHis 736
              |||  ||||||:::  |||  :::  ||||||:::||||||
Db       2185 CCCCCATTCCTCTGGAGAATTAAACCATTCACTGCCTGCCAGGAAGTCCCGTTCTCCCGC 2244

Qy        737 IleTrpAsnGlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerPro 756
              :::|||  ::::::  |||||||||  |||:::|||||||||  :::|||
Db       2245 GTGTGGTGCAGTAACCGGCAGCCCCTGCACTGTGCCTTCTCCCTGGAGCGTTATACGCCC 2304

Qy        757 SerThrSerAspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSer 776
              :::||||:::  |||:::||||||:::  :::  |||::::::|||  |||
Db       2305 ACTACCACCCAGCTGTCCTGCAAAATCTGCATTCGGCAGCTCAAAGGCCATGAACAGATC 2364

Qy        777 PheSerIleAsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGlu 796
              :::  :::|||  :::  |||  :::
Db       2365 CTCCAAGTGCAGACATCAATCCTAGAGAGTGAACGAGAAACCATCACTTTCTTCGCACAA 2424

Qy        797 SerGluAlaGlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIle 816
              ::::::  ||||||  ||||||  |||||||||||||||:::  |||
Db       2425 GAGGACAGCACTTTCCCTGCACAGACTGGCCCCAAAGCCTTCAAAATTCCCTACTCCATC 2484

Qy        817 ArgGlnLysIleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThr 836
              ||||||:::|||  ::::::  |||  |||  :::|||  ||||||:::
Db       2485 AGACAGCGGATTTGTGCTACATTTGATACCCCCAATGCCAAAGGCAAGGACTGGCAGATG 2544

Qy        837 LeuAlaGlnLysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerPro 856
              ||||||||||||  ::::::  :::||||||:::||||||::::::  ||||||
Db       2545 TTAGCACAGAAAAACAGCATCAACAGGAATTTATCTTATTTCGCTACACAAAGTAGCCCA 2604

Qy        857 ThrAlaMetIleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGln 876
              :::|||:::||||||||||||||||||||||||  :::|||:::|||
Db       2605 TCTGCTGTCATTTTGAACCTGTGGGAAGCTCGTCATCAGCATGATGGTGATCTTGACTCC 2664

Qy        877 LeuAlaAlaAlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGlu 896
              ||||||  |||:::  :::|||:::  |||  :::||||||
Db       2665 CTGGCCTGTGCCCTTGAAGAGATTGGGAGGACACACACGAAACTCTCAAACATTTCAGAA 2724

Qy        897 AlaGlu 898
              ::::::
Db       2725 TCCCAG 2730

RESULT 10
AY411748
LOCUS       AY411748                2532 bp    DNA     linear   GSS 12-DEC-2003
DEFINITION  Pan troglodytes HCM4327 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY411748
VERSION     AY411748.1  GI:39767716
KEYWORDS    GSS.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1  (bases 1 to 2532)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 2532)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1..2532
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
     gene            <1..>2532
                     /locus_tag="HCM4327"
ORIGIN

Alignment Scores:
Pred. No.:              3.76e-157     Length:          2532
Score:                  1890.00       Matches:         390
Percent Similarity:     60.87%        Conservative:    125
Best Local Similarity:  46.10%        Mismatches:      275
Query Match:            39.48%        Indels:          56
DB:                     29            Gaps:            16

US-09-970-944-2 (1-899) x AY411748 (1-2532)

Qy        103 GluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGluGluTyr 122
              ||||||:::|||  ||||||||||||||||||||||||::::::|||||||||||:::|||
Db         13 GAGGTGCAGATCGAGGTGTCGCGGCAGCAGGTGGAGGAGCTCTTTGGGCTGGAGGATTAC 72

Qy        123 TrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAlaTyrIle 142
              ||||||||||||||||||||||||:::||||||||||||||||||::::::||||||:::
Db         73 TGGTGCCAGTGCGTGGCCTGGAGCTCTGCGGGCACCACCAAGAGTCGCCGAGCCTACGTC 132

Qy        143 ArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluValSerLeu 162
              ||||||||
Db        133 CGCATCGCCTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 192

Qy        163 GluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGluValGlu 182
                                                                    |||||
Db        193 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNGTGGAA 252

Qy        183 TrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIleThrArg 202
              ||||||:::||||||||||||::::::||||||:::  |||  |||  :::|||
Db        253 TGGCTCAAGAATGAGGATGTCATCGACCCCACCCAGGACACCAACTTCCTGCTCACCATC 312

Qy        203 GluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThrCysVal 222
              :::|||:::|||::::::||||||||||||:::||||||||||||||||||||||||
Db        313 GACCACAACCTCATCATCCGCCAGGCCCGCCTGTCGGACACTGCCAACTATACCTGCGTG 372

Qy        223 AlaLysAsnIleValAlaArgArgArgSerAlaSerAlaAlaValIleValTyrValAsn 242
              ||||||||||||||||||:::||||||||  :::|||  |||||||||||||||||||
Db        373 GCCAAGAACATCGTGGCCAAACGCCGGAGCACCACTGCCACCGTCATCGTCTACGTGAAT 432

Qy        243 GlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArgGlyTrp 262
              ||||||||||||:::|||  |||||||||  ||||||  ||||||  |||
```

```
Db    433 GGCGGCTGGTCCAGCTGGGCAGAGTGGTCGCCCTGCTCCAACCGCTGTGGNNGANNCTGG 492

Qy    263 GlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPheCysGlu 282
Db    493 CAGAAGCGCACCCGGACCTGCACCAACCCCGCCCCACTCAACGGAGGGNNNNTCTGCGAG 552

Qy    283 GlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGlySerTrp 302
Db    553 GGC---CANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNTCGATGGGGCGTGG 609

Qy    303 SerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSerArgGlu 322
Db    610 ACGGAGTGGAGCAAGTGGTCAGCCTGCAGCACTGAGTGTGCCCACTGGCGTAGCCGCGAG 669

Qy    323 CysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeuAspThr 342
Db    670 TGCATGGCGCCCCCACCCCAGAACGGAGGCCGTGACTGCAGCGGGACGCTGCTCGACTCT 729

Qy    343 ArgAsnCysThrSerAspLeuCysValHisSer--------------------------- 353
Db    730 AAGAACTGCACAGATGGGCTGTGCATGCAAAATAAGAAAACTCTAAGCGACCCCAACAGC 789

Qy    354 ------------AlaSerGlyProGluAspValAlaLeuTyrValGlyLeu---IleAla 368
Db    790 CACCTGCTGGAGGCCTCAGGG------GATGCGGCGCTGTATGCGGGGCTCGTGGTGGCC 843

Qy    369 ValAlaValCysLeuValLeuLeuLeuLeuValLeuIleLeuValTyrCysArgLysLys 388
Db    844 GTCTTCNTGGTCNTGGCAATCCTCATGGCGGTGGGGGTGGTGGTGTACCGCCGCAACTGC 903

Qy    389 GluGlyLeuAspSerAspValAlaAspSerSer---IleLeuThrSerGlyPheGlnPro 407
Db    904 CGTGACTTCGACACAGACATCACTGACTCATCTGCTGCCCTGACTGGTGGTTTCCACCCC 963

Qy    408 ValSerIleLysProSerLysAlaAspAsnProHisLeuLeu------ThrIleGlnPro 425
Db    964 GTCAACTTTAAGACGGCAAGGCCCAGCAACCCGCAGCTCCTACACCCCTCTGTGCCTCCT 1023

Qy    426 AspLeuSerThrThrThr---ThrTyrGlnGlySerLeuCysProArgGlnAspGlyPro 444
Db   1024 GACCTGACAGCCAGCGCCGGCATCTACCGCGGACCCGTGTATGCCCTGCAGGACTCC--- 1080

Qy    445 SerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeuGlyGlyGlyArgHis 464
Db   1081 ACCGACAAAATCCCCATGACCAACTCTCCTCTGCTGGACCCCTTACCCAGCCTTAAGGTC 1140

Qy    465 ThrLeuHisHisSerSerProThr------------------SerGluAlaGluGluPhe 478
Db   1141 AAGGTCTACAGCTCCAGCACTACGGGCTCTGGGCCAGGCCTGGCAGATGGGGCTGACCTG 1200

Qy    479 ValSerArgLeuSerThrGlnAsnTyr--------------------------------- 487
Db   1201 CTGGGAGTCTTGCCGCCTGGCACATACCCTAGCGATTTCGCCCGGGACACCCACTTCCTG 1260

Qy    488 ---PheArgSer------------------------------LeuProArgGlyThrSer 496
Db   1261 CACCTGCGCAGCGCCAGCCTCGGTTCCCAGCAGCTCTTGGGCCTGCCCCGAGACCCAGNN 1320

Qy    497 AsnMetThrTyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGly 516
Db   1321 AGCAGCGTCAGCGGCACCTTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNGG 1380

Qy    517 IleSerLeuLeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeu 536
Db   1381 GTCAGCTTGCTGGTGCCCAATGGAGCCATTCCCCAGGGCAAGTTCTACGAGATGTATCTA 1440

Qy    537 ThrLeuHisLysProGluAspValArgLeuProLeuAla---GlyCysGlnThrLeuLeu 555
Db   1441 CTCATCAACAAGGCAGAAAGTACC---CTCCCGCTTTCAGAAGGGACCCAGACAGTATTG 1497

Qy    556 SerProIleValSerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAla 575
Db   1498 AGCCCCTCAGTGACCTGTGGACCCACAGGCCTCCTGCTGTGCCACCCCGTCATCCTCACC 1557

Qy    576 MetAspHisCysGlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSer 595
Db   1558 ATGCCCCACTGTGCCGAAGTCAGTGCCCGTGACTGGATCTTTCAGCTCAAGACCCAGGCC 1617

Qy    596 CysGluGlySerTrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeu 615
Db   1618 CACCAGGGCCACTGGGAG---GAGGTGGTGACCCTGGATGAGGAGACCCTGAACACACCC 1674

Qy    616 TyrTyrCysGlnLeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPhe 635
Db   1675 TGCTACTGCCAGCTGGAGCCCAGGGCCTGTCACATCCTGCTGGACCAGCTGGGCACCTAC 1734

Qy    636 AlaLeuValGlyGluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPhe 655
Db   1735 GTGTTCACGGGCGAGTCCTATTCCCNNNNNNCAGTCNNNNNGCTCCNGCTGGCCGTCTTN 1794

Qy    656 AlaProValAlaCysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThr 675
Db   1795 GCCCCTGCCCTCTGCACCTCCCTGGANNNNNGCCTCCGGGTCTACTGCCTGGNNNNNNNN 1854

Qy    676 HisAspAlaLeuLysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGln 695
Db   1855 NNNNNNNNNNNNNNNGAGGTGCTGGAGCTGGAGCGGACTCTGGGCGGATACTTGGTGGAG 1914

Qy    696 GluProArgValLeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAsp 715
Db   1915 GAGCCGAAACCGCTAATGTTCAAGGACAGTTACCACAACCTGCGCCTCTCCCTCCATGAC 1974

Qy    716 ValProSerSerLeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyr 735
Db   1975 CTCCCCCATGCCCATTGGAGGAGCAAGCTGCTGGCCAAATACCAGGAGATCCCCTTCTGT 2034

Qy    736 HisIleTrpAsnGlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSer 755
Db   2035 CACATTTGGAGTGGCAGCCAGAAGGCCCTCCACTGCACTTTCACCCTGGAGAGGCACAGC 2094

Qy    756 ProSerThrSerAspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGln 775
Db   2095 TTGGCCTCCACAGAGCTCACCTGCAAGATCTGCGTGCGGCAAGTGGAAGGGGAGGGCCAG 2154

Qy    776 SerPheSerIleAsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeu 795
Db   2155 ATATTCCAGCTGCATACCACTCTGGCA---GAGACACCTGCTGGCTCCCTGGACACTCTC 2211

Qy    796 GluSerGluAlaGly------ValProAlaLeuValGlyProSerAlaPheLysIlePro 813
Db   2212 TGCTCTGCCCCTGGCAGCACTGTCACCACCCAGCTGGGACCTTATGCCTTCAAGATCCCA 2271

Qy    814 PheLeuIleArgGlnLysIleIleSerSerLeuAspProProCysArgArgGlyAlaAsp 833
Db   2272 CTGTCCATCCGCCAGAAGATATGCAACAGCCTAGATGCCCCCAACTCACGGGGCAATGAC 2331

Qy    834 TrpArgThrLeuAlaGlnLysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLys 853
Db   2332 TGGCGGATGTTAGCACAGAAGCTCTCTATGGACCGGTACCTGAATTACTTTGCCACCAAA 2391

Qy    854 ProSerProThrAlaMetIleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsn 873
Db   2392 GCGAGCCCCACGGGTGTGATCCTGGACCTCTGGGAAGCTCTGCAGCAGGACGATGGGGAC 2451

Qy    874 LeuSerGlnLeuAlaAlaAlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThr 893
Db   2452 CTCAACAGCCTGGCGAGTGCCTTGGAGGAGATGGGCAAGAGTGAGATGCTGGTGGCTGTG 2511

Qy    894 ValSerGluAlaGluCys 899
Db   2512 GCCACCGACGGGGACTGC 2529
```

RESULT 11
BC033727
LOCUS BC033727 1532 bp mRNA linear HTC 19-NOV-2003
DEFINITION Homo sapiens netrin receptor Unc5h1, mRNA (cDNA clone

```
            IMAGE:5166762), containing frame-shift errors.
ACCESSION   BC033727
VERSION     BC033727.1  GI:21707230
KEYWORDS    HTC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1532)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  MEDLINE   22388257
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1532)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUL-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC),
            Gaithersburg, Maryland;
            Web site: http://www.nisc.nih.gov/
            Contact: nisc_mgc@nhgri.nih.gov
            Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
            Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
            Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
            Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
            Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D.,McCloskey,J.C.,
            McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
            Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
            Young,A., Zhang,L.-H. and Green,E.D.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 68 Row: i Column: 2
            This clone was selected for full length sequencing because it
            passed the following selection criteria: GenomeScan gene prediction
            This clone has the following problem: frame shifted.
FEATURES             Location/Qualifiers
     source          1..1532
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5166762"
                     /tissue_type="Brain, adult medulla"
                     /clone_lib="NIH_MGC_119"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
ORIGIN

Alignment Scores:
Pred. No.:               2.91e-155     Length:          1532
Score:                   1864.50       Matches:         377
Percent Similarity:      67.08%        Conservative:    2
Best Local Similarity:   66.73%        Mismatches:      2
Query Match:             38.95%        Indels:          184
DB:                      11            Gaps:            3

US-09-970-944-2 (1-899) x BC033727 (1-1532)

Qy      23 SerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnProAspLeu 42
           :::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11 GCAGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTG 70

Qy      43 LeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysProValLeu 62
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      71 CTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTG 130

Qy      63 LeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGluTrpVal 82
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     131 CTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTG 190

Qy      83 ArgGlnValAspHisValIleGluArgSerThrAspGlySerSer--------------- 97
           |||||||||||||||||||||||||||||||||||||||||||||
Db     191 CGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGAGGGAGATAGAGAG 250

Qy      97 ------------------------------------------------------------ 97

Db     251 GGTTGGCGGCCAATGTAGCCCAACAGAACCCTGGGCACCAACACCTCCCAGGACCAGAGC 310

Qy      97 ------------------------------------------------------------ 97

Db     311 CTCTCCCGGCCCCAGCGGAGGTGCCAGAGGCCTCCCTGCCCAGCAGCAGGGCCCACCTTG 370

Qy      97 ------------------------------------------------------------ 97

Db     371 TCTTGCTGGGAGGGTCTGCAGGAGAACCGCAGGGTCCTCTTCTGCCCCGAGCATCCGACT 430

Qy      97 ------------------------------------------------------------ 97

Db     431 CCAGCCTCCCATGGCCGGAACATCTGAACGTTCATTGATCTCATCGATTTCTCAACCCAA 490

Qy      97 ------------------------------------------------------------ 97

Db     491 AGCCATCGAGTGCTTTGAGGTGAAGAAAAAGGCTTTCCTTACCCACGGCAGATACCACGG 550

Qy      97 ------------------------------------------------------------ 97

Db     551 CAGTGGCGCCACGCCGCCAAAGACCAAAGACCCCAAACCAGAAACGTTCTGTGGTCAGAC 610

Qy      98 -GlyGluProThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPh 117
            |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     611 AGGGCTGCCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTT 670

Qy     117 eGlyLeuGluGluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLysSe 137
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     671 CGGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAG 730

Qy     137 rGlnLysAlaTyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAl 157
           ||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
Db     731 TCAGAAGGCCTACATCCGCATAGCCTATTTGCGCAAGAACTTCGAGCAGGAGCCGCTGGC 790

Qy     157 aLysGluValSerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIlePr 177
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     791 CAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCC 850

Qy     177 oProAlaGluValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAs 197
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     851 TCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAA 910
```

```
Qy      197 nValTyrIleThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAl 217
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      911 TGTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGC 970

Qy      217 aAsnTyrThrCysValAlaLysAsnIleValAlaArgArgArgSerAlaSerAlaAlaVa 237
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      971 CAACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGT 1030

Qy      237 lIleValTyrValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSe 257
            ||||||||||
Db     1031 CATCGTCTAC-------------------------------------------------- 1040

Qy      257 rCysGlyArgGlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGl 277

Db     1040 ------------------------------------------------------------ 1040

Qy      277 yGlyAlaPheCysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSe 297

Db     1040 ------------------------------------------------------------ 1040

Qy      297 rValAspGlySerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHi 317
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1041 -GTGGACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCA 1099

Qy      317 sTrpArgSerArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGl 337
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1100 CTGGCGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGG 1159

Qy      337 yThrAspLeuAspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyPr 357
            |||||||||||||||||||||||||||||||||||||||||||:::||||||||||||||
Db     1160 CACTGACCTGGACACCCGCAACTGTACCAGTGACCTCTGTGTACACACTGCTTCTGGCCC 1219

Qy      357 oGluAspValAlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLe 377
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1220 TGAGGACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCT 1279

Qy      377 uLeuValLeuIleLeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAs 397
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1280 GCTTGTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGA 1339

Qy      397 pSerSerIleLeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAs 417
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1340 CTCGTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAA 1399

Qy      417 nProHisLeuLeuThrIleGlnProAspLeuSer---ThrThrThrThrTyrGlnGlySe 436
            ||||||||||||||||||||||||||||||||||   |||||||||||||||||||||||
Db     1400 CCCCCATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAG 1459

Qy      436 rLeuCysProArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLe 456
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1460 TCTCTGTCCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCT 1519

Qy      456 uSerProLeuGly 460
            |||||||||||||
Db     1520 CAGCCCCCTGGGT 1532
```

RESULT 12
AY401470
LOCUS       AY401470                2507 bp    DNA     linear   GSS 12-DEC-2003
DEFINITION  Pan troglodytes HCM0901 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
ACCESSION   AY401470
VERSION     AY401470.1  GI:39757459
KEYWORDS    GSS.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1  (bases 1 to 2507)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 2507)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES             Location/Qualifiers
     source          1..2507
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
     gene            <1..>2507
                     /locus_tag="HCM0901"
ORIGIN

Alignment Scores:

| | | | |
|---|---|---|---|
| Pred. No.: | 9.18e-127 | Length: | 2507 |
| Score: | 1550.00 | Matches: | 332 |
| Percent Similarity: | 52.59% | Conservative: | 135 |
| Best Local Similarity: | 37.39% | Mismatches: | 257 |
| Query Match: | 32.38% | Indels: | 164 |
| DB: | 29 | Gaps: | 14 |

US-09-970-944-2 (1-899) x AY401470 (1-2507)

```
Qy      103 GluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGluGluTyr 122
            |||||    ||||||||:::||||||||||||            |||     |||:::|||
Db       15 GAAGTGTTCATCAATGTTACCAGGCAACAGGTGGAGGACTTCCATGGGCCCGAGGACTAT 74

Qy      123 TrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAlaTyrIle 142
            |||||||||||||||||||||         ||||||:::||||||:::||||||    :::
Db       75 TGGTGCCAGTGTGTGGCGTGGAGCCACCTGGGTACCTCCAAGAGCAGGAAGGCCTCTGTG 134

Qy      143 ArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluValSerLeu 162
            |||||||||   ||||||||||||||||||||:::|||         :::||||||   :::
Db      135 CGCATAGCCTATTTACGGAAAAACTTTGAACAAGACCCACAAGGAAGGGAGGTTCCCATT 194

Qy      163 GluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGluValGlu 182
            |||      |||||||||   ||||||||||||||||:::|||     |||   ||||||
Db      195 GAAGGCATGATTGTACTGCACTGCCGCCCACCAGAGGGAGTCCCTGCTGCNNNNGTGGAA 254

Qy      183 TrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIleThrArg 202
            ||||||:::||||||:::      :::|||            |||   |||:::
Db      255 TGGCTGAAAAATGAAGAGCCCATTGACTCTGAACAAGACGAGAACATTGACACCAGGGCT 314

Qy      203 GluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThrCysVal 222
            :::|||:::|||::::::||||||||||||||:::|||:::      |||||||||||:::
Db      315 GACCATAACCTGATCATCAGGCAGGCACGGCTCTCGGACTCAGGAAATTACACCTGCATG 374

Qy      223 AlaLysAsnIleValAlaArgArgArgSerAlaSerAlaAlaValIleValTyrValAsn 242
            |||   |||||||||||:::|||||||||      ||||||   |||:::||||||||||||
Db      375 GCAGCCAACATCGTGGCTAAGAGGAGAAGCCTGTCGGCCACTGTTGTGGTCTACGTGAAT 434

Qy      243 GlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArgGlyTrp 262
            ||||||||||||:::||||||||||||||||   |||:::         ||||||||||||
Db      435 GGAGGCTGGTCTTCCTGGACAGAGTGGTCAGCCTGCAATGTTCGCTGTGGTAGAGGATGG 494

Qy      263 GlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPheCysGlu 282
            |||||||||||||||:::|||||||||||||||||||||||||||||||||||||||||||
Db      495 CAGAAACGTTCCCGGACCTGCACCAACCCAGCTCCTCTCAATGGTGGGGCCTTTTGTGAG 554

Qy      283 GlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGlySerTrp 302
            |||   :::|||         |||
Db      555 GGAATGTCAGTGCAGAAAATAACC------------------------------------ 578
```

```
Qy        303 SerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSerArgGlu 322
Db        578 ------------------------------------------------------------ 578
Qy        323 CysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeuAspThr 342
Db        578 ------------------------------------------------------------ 578
Qy        343 ArgAsnCysThrSerAspLeuCys------------------------------------ 350
Db        579 ------TGCACTTCT---CTTTGTCCTGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 629
Qy        350 ------------------------------------------------------------ 350
Db        630 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 689
Qy        350 ------------------------------------------------------------ 350
Db        690 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 749
Qy        351 ---------------ValHisSerAlaSerGlyProGluAspValAlaLeuTyrValGly 365
Db        750 NNNNNNNNNNNNNGCATTGAGAATGCCAGC---------GACATTGCTTTGTACTCGGGC 800
Qy        366 LeuIleAlaValAlaValCysLeuValLeuLeuLeuLeuValLeuIleLeuValTyrCys 385
Db        801 TTG---GGCGCTGCCGTCGTGGCCGTTGCAGTCCTGGTCATTGGTGTCACCCTTTACAGA 857
Qy        386 ArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIleLeuThrSerGlyPhe 405
Db        858 CGGAGCCAGAGTGACTATGGCGTGGACGTCATTGACTCTTCTGCATTGACAGGTGGCTTC 917
Qy        406 GlnProValSerIleLysProSerLysAlaAspAsnProHisLeuLeu------ThrIle 423
Db        918 CAGACCTTCAACTTCAAAACAGTCCGTCAAGNNACTCCCTGCTCCTGAATTCCGCCATG 977
Qy        424 GlnProAspLeuSerThrThrThrThrTyrGlnGlySerLeuCysProArgGlnAspGly 443
Db        978 CAGCCAGATCTGACAGTGAGCCGGACATACAGCGGACCCATCTGT---CTGCAGGAC--- 1031
Qy        444 ProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeuGly--------- 460
Db       1032 CCTCTGGACAAGGAGCTCATGACAGAGTCCTCACTGTTTAACCCTTTGTCGGACATCAAA 1091
Qy        460 ------------------------------------------------------------ 460
Db       1092 GTCAAAGTCCAGAGCTCGTTCATGGTTTCCCTGGGAGTGTCTGAGAGAGCTGAGTACCAC 1151
Qy        461 ------------------------------GlyGlyArgHisThrLeuHisHisSerSer 470
Db       1152 GGCAAGAATCATTCCAGGACTTTTCCCCATGGAAACAACCACAGCTTTAGTACAATGCAT 1211
Qy        471 ProThrSerGluAlaGluGluPheValSerArgLeuSerThrGlnAsnTyrPheArgSer 490
Db       1212 CCCAGAAATAAAATG---CCCTACATCCAAAATCTGTCA-----------------TCA 1250
Qy        491 LeuProArgGlyThrSerAsnMetThrTyrGlyThrPheAsnPheLeuGlyGlyArgLeu 510
Db       1251 CTCCCCACAAGGACAGAACTGAGGACAACTGGTGTCTNNGCCATTTAGGGGGGCGNNNN 1310
Qy        511 MetIleProAsnThrGlyIleSerLeuLeuIleProProAspAlaIleProArgGlyLys 530
Db       1311 NNNNNN---NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1367
Qy        531 IleTyrGluIleTyrLeuThrLeuHisLysProGluAspValArgLeuProLeuAlaGly 550
Db       1368 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNCCTCCAGTCAGATGGC 1427
Qy        551 CysGlnThrLeuLeuSerProIleValSerCysGlyProProGlyValLeuLeuThrArg 570
Db       1428 TCTGAGGTGCTCCTGAGTCCTGAAGTCACCTGTGGTCCTCCAGACATGATCGTCACCACT 1487

Qy        571 ProValIleLeuAlaMetAspHisCysGlyGluProSerProAspSerTrpSerLeuArg 590
Db       1488 CCCTTTGCATTGACCATCCCACACTGTGCAGATGTCAGTNNNGAGCACTGGAATATCCAT 1547
Qy        591 LeuLysLysGlnSerCysGluGlySerTrpGluGlnAspValLeuHisLeuGlyGluGlu 610
Db       1548 TTAAAGAAGAGGACACAGCAGGGCAAATGGGAG---GAAGTGATGTCAGTGGAAGATGAA 1604
Qy        611 AlaProSerHisLeuTyrTyrCysGlnLeuGluAlaSerAlaCysTyrValPheThrGlu 630
Db       1605 TCTACATCC------TGTTACTGCCTTTTGGACCCCTTTGCATGTCATGTGCTCCTGGAC 1658
Qy        631 GlnLeuGlyArgPheAlaLeuValGlyGluAlaLeuSerValAlaAlaAlaLysArgLeu 650
Db       1659 AGCTTTGGGACCTATGCGCTCACTGGAGAGCCAATCACAGACTGTGCCGTGAAGCAACTG 1718
Qy        651 LysLeuLeuLeuPheAlaProValAlaCysThrSerLeuGluTyrAsnIleArgValTyr 670
Db       1719 AAGGTGGCGGTTTTTGGCTGCATGTCCTGTAACTCCCTGGATTACAACTTGAGAGTTTAC 1778
Qy        671 CysLeuHisAspThrHisAspAlaLeuLysGluValValGlnLeuGluLysGlnLeuGly 690
Db       1779 TGTGTGGACAATACCCCTTGTGCATTTCANGAAGTGGTTTCAGATGAAAGGCATCAAGGT 1838
Qy        691 GlyGlnLeuIleGlnGluProArgValLeuHisPheLysAspSerTyrHisAsnLeuArg 710
Db       1839 GGACAGCTCCTGGAAGAACCAAAATTGCTGCATTTCAAAGGGAATACCTTTAGTCTTCAG 1898
Qy        711 LeuSerIleHisAspValProSerSerLeuTrpLysSerLysLeuLeuValSerTyrGln 730
Db       1899 ATTTCTGTCCTTGATATTCCCCCATTCCTCTGGAGAATTAAACCATTCACTGCCTGCCAG 1958
Qy        731 GluIleProPheTyrHisIleTrpAsnGlyThrGlnArgTyrLeuHisCysThrPheThr 750
Db       1959 GAAGTCCCGTTCTCCCGCGTGTGGTGCAGTAACCGGCAGCCCCTGCACTGTGCCTTCTCC 2018
Qy        751 LeuGluArgValSerProSerThrSerAspLeuAlaCysLysLeuTrpValTrpGlnVal 770
Db       2019 CTGGAGCGTTATACGCCCACTACCACCCAGCTGTCCTGCAAAATCTGCATTCGGCAGCTC 2078
Qy        771 GluGlyAspGlyGlnSerPheSerIleAsnPheAsnIleThrLysAspThrArgPheAla 790
Db       2079 AAAGGCCATGAACAGATCCTCCAAGTGCAGACATCAATCCTAGAGAGTGAACGAGAAACC 2138
Qy        791 GluLeuLeuAlaLeuGluSerGluAlaGlyValProAlaLeuValGlyProSerAlaPhe 810
Db       2139 ATCACTTTCTTTGCACAAGAGGACAGCACTTTCCCTGCACAGACTGGCCCCAAAGCCTTC 2198
Qy        811 LysIleProPheLeuIleArgGlnLysIleIleSerSerLeuAspProProCysArgArg 830
Db       2199 AAAATTCCCTACTCCATCAGACAGCGGATTTGTGCTACATTTGATACCCCCAATGCCAAA 2258
Qy        831 GlyAlaAspTrpArgThrLeuAlaGlnLysLeuHisLeuAspSerHisLeuSerPhePhe 850
Db       2259 GGCAAGGACTGGCAGATGTTAGCACAGAAAAACAGCATCAACAGGAATTTATCTTATTTC 2318
Qy        851 AlaSerLysProSerProThrAlaMetIleLeuAsnLeuTrpGluAlaArgHisPhePro 870
Db       2319 GCTACACAAAGTAGCCCATCTGCTGTCATTTTGAACCTGTGGGAAGCTCGTCATCAGCAT 2378
Qy        871 AsnGlyAsnLeuSerGlnLeuAlaAlaAlaValAlaGlyLeuGlyGlnProAspAlaGly 890
Db       2379 GATGGTGATCTTGACTCCCTGGCCTGTGCCCTTGAAGAGATTGGGAGGACACACACGAAA 2438
Qy        891 LeuPheThrValSerGluAlaGlu 898
Db       2439 CTCTCAAACATTTCAGAATCCCAG 2462

RESULT 13
BI758231
LOCUS       BI758231              1034 bp    mRNA    linear   EST 25-SEP-2001
DEFINITION  603029876F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5200171 5',
            mRNA sequence.
```

```
ACCESSION   BI758231
VERSION     BI758231.1  GI:15749809
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1034)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11501  row: g  column: 20
            High quality sequence stop: 793.
FEATURES             Location/Qualifiers
     source          1. .1034
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5200171"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_114"
                     /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: EcoRV (destroyed); RNA source anonymous pool of 6
                     male brains, age range 23-27 yo. Library is oligo-dT
                     primed and directionally cloned (EcoRV site is  destroyed
                     upon cloning).  Average insert size 1.5 kb, insert size
                     range 1-3 kb.  Library is normalized and enriched for
                     full-length  clones and was constructed by C. Gruber
                     (Invitrogen).  Research  Genetics tracking code 019. Note:
                     this is a NIH_MGC Library."
ORIGIN

Alignment Scores:
Pred. No.:              1.13e-111     Length:         1034
Score:                  1373.50       Matches:        291
Percent Similarity:     88.82%        Conservative:   11
Best Local Similarity:  85.59%        Mismatches:     28
Query Match:            28.69%        Indels:         14
DB:                     12            Gaps:           2

US-09-970-944-2 (1-899) x BI758231 (1-1034)

Qy         36 ProGlyAlaAsnProAspLeuLeuProHisPheLeuValGluProGluAspValTyrIle 55
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 CCTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATC 61

Qy         56 ValLysAsnLysProValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePhe 75
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 GTCAAGAACAAGCCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTC 121

Qy         76 LysCysAsnGlyGluTrpValArgGlnValAspHisValIleGluArgSerThrAspGly 95
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 AAGTGCAACGGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGG 181

Qy         96 SerSerGlyGluProThrMetGluValArgIleAsnValSerArgGlnGlnValGluLys 115
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 AGCAGTGGGCTGCCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAG 241

Qy        116 ValPheGlyLeuGluGluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThr 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 GTGTTCGGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACC 301

Qy        136 LysSerGlnLysAlaTyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluPro 155
              ||||||||||||||||||||||||||||    ||||||||||||||||||||||||||||
Db        302 AAGAGTCAGAAGGCCTACATCCGCATAGCCTATTTGCGCAAGAACTTCGAGCAGGAGCCG 361

Qy        156 LeuAlaLysGluValSerLeuGluGlnGlyIleValLeuProCysArgProProGluGly 175
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 CTGGCCAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGC 421

Qy        176 IleProProAlaGluValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAsp 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 ATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGAC 481

Qy        196 ProAsnValTyrIleThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAsp 215
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 CCCAATGTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGAC 541

Qy        216 ThrAlaAsnTyrThrCysValAlaLysAsnIleValAlaArgArgArgSerAlaSerAla 235
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 ACGGACAACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCT 601

Qy        236 AlaValIleValTyrValAsnGlyGlyTrpSerThr-TrpThrGluTrpSerValCysSe 255
              |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db        602 GCTGTCATCGTCTACGTGAACGGTGGGTGGTCGACCGTGGACCGAGTGGTCCGTCTGCGC 661

Qy        255 rAlaSerCysGlyArgGlyTrpGlnLys-ArgSerArgSerCysThrAsnProAlaProL 275
               ||||||||||||||||||||||||||| ||||||||||||||||     ||||||    ||||
Db        662 -GCCAGCTGTGGGCGCGGCTGGCAGAAACCGGAGCCGGAGCTGCAC-AACCCGGTGCCTC 719

Qy        275 euAsnGlyGlyAlaPheCysGlu-GlyGlnAsnValHisAspArgThrValSerSerLeu 294
              ||||||||||||||||||||||| ||||||||||||           :::     ::::::
Db        720 TCAACGGGGGCGCTTTCTGTGAGGGGGCAGAATGTC------CAGAAAGCAGCTGCGCCA 773

Qy        295 LeuValSerValAspGlySerTrpSerProTrpSerLysTrpSerAlaCys--GlyLeuA 314
                 |||    |||||||||||    |||    ||||||:::||||||||||||  ||||||
Db        774 CCTGTGCCAGTGGACGGCAGCTGTAGCACGTGGAGCCAGTGGTCGGCCTGTTGGGCTTGG 833

Qy        314 spCysThrHis-TrpArgSerArgGluCysSerAsp-ProAlaProArgAsnGlyGlyGl 333
                ||||||||  |||          ||||||||::: |||||||||||:::|||
Db        834 CTTGCACCCACTTGGGCGGAGCCGGAGTGCTCTGAACCCAGCACCCCGGCACGGCAGGGG 893

Qy        333 uGluCys---GlnGlyThrAspLeuAspThrArgAsnCysThrSerAspLeuCysValHi 352
                  |||     |||    |||          ||||||:::     :::|||:::||||||||
Db        894 GAGGTGTGCCCAGGTCACTGGACCTGGCACCCGGGATGG-TCCAGTGAGCTCTGTGTCCC 952

Qy        352 sSerAlaSerGlyProGluAspValAlaLeuTyrValGlyLeuIle 367
                :::    ||||||||||||||:::|||    ||||||||||||||
Db        953 -ACGGGTTCTGGCCCTGAGGACTTGGCCTCCTATGTGGGCCTCATC 997

RESULT 14
BI818609
LOCUS       BI818609                 818 bp    mRNA    linear   EST 04-OCT-2001
DEFINITION  603033362F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5174559 5',
            mRNA sequence.
ACCESSION   BI818609
VERSION     BI818609.1  GI:15929902
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 818)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11434  row: l  column: 16
            High quality sequence stop: 744.
```

```
FEATURES             Location/Qualifiers
     source          1..818
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5174559"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_115"
                     /note="Organ: pooled brain, lung, testis; Vector:
                     pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
                     source anonymous pool of 6 male brains, age range 23-27; 1
                     male lung, age 27; and 1 male testis, age 69.  Library is
                     oligo-dT primed and directionally cloned (EcoRV site is
                     destroyed upon cloning).  Average insert size 1.8 kb,
                     insert size range 1-3 kb. Library is normalized and
                     enriched for full-length clones and was constructed by C.
                     Gruber (Invitrogen).  Research Genetics tracking code
                     021. Note: this is a NIH_MGC Library."
ORIGIN

Alignment Scores:
Pred. No.:               5.39e-109       Length:          818
Score:                   1341.50         Matches:         263
Percent Similarity:      80.61%          Conservative:    3
Best Local Similarity:   79.70%          Mismatches:      5
Query Match:             28.02%          Indels:          59
DB:                      12              Gaps:            1

US-09-970-944-2 (1-899) x BI818609 (1-818)

Qy          1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 63

Qy         21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         64 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 123

Qy         41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 183

Qy         61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 243

Qy         81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        244 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 303

Qy        101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 363

Qy        121 GluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 423

Qy        141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160
              ||||||||||||||    ||||||||||||||||||||||||||||||||||||||||||
Db        424 TACATCCGCATAGCCTATTTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 483

Qy        161 SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 543

Qy        181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        544 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 603

Qy        201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        604 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 663

Qy        221 CysValAlaLysAsnIleValAlaArgArg-ArgSerAlaSerAlaAlaValIleValTy 240
              ||||||||||||||||||||||||||:::  |||    ||||||||||||||||||||||
Db        664 TGCGTGGCCAAGAACATCGTGGCACGTCAGCCGCAGGGCCTCCGCTGCTGTCATCGTCTA 723

Qy        240 rValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyAr 260
              |
Db        724 C----------------------------------------------------------- 724

Qy        260 gGlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPh 280

Db        724 ------------------------------------------------------------ 724

Qy        280 eCysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGl 300
                                                                  ||||||||
Db        725 ----------------------------------------------------GTGGACGG 732

Qy        300 ySerTrpSerPro-TrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgS 320
              ||||||||||||| ||||||:::||||||    ||||||||||||||||||||||||||
Db        733 CAGCTGGAGCCCGTTGGAGCCAGTGGTCGGGCTGTGGGCTGGACTGCACCCACTGGCGGA 792

Qy        320 erArgGluCysSerAspProAlaPro 328
              |||||:::||||||||||||    |||
Db        793 GCCGGAAGTGCTCTGACCCAAACCCG 818
```

```
RESULT 15
BQ689148
LOCUS       BQ689148                874 bp    mRNA    linear   EST 15-JUL-2002
DEFINITION  AGENCOURT_8064396 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6206234
            5', mRNA sequence.
ACCESSION   BQ689148
VERSION     BQ689148.1  GI:21814464
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 874)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2360  row: g  column: 03
            High quality sequence stop: 737.
FEATURES             Location/Qualifiers
     source          1..874
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6206234"
                     /tissue_type="ductal carcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_110"
                     /note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G). Library constructed by
                     Ling Hong in the laboratory of Gerald M. Rubin (University
                     of California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library."
ORIGIN

Alignment Scores:
Pred. No.:               6.57e-103       Length:          874
```

```
Score:                    1274.00      Matches:        255
Percent Similarity:       63.59%       Conservative:   0
Best Local Similarity:    63.59%       Mismatches:     0
Query Match:              26.61%       Indels:         146
DB:                       13           Gaps:           1

US-09-970-944-2 (1-899) x BQ689148 (1-874)

Qy        499 ThrTyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSer 518
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGAATCAGC 60

Qy        519 LeuLeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeu 538
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTCCTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTG 120

Qy        539 HisLysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIle 558
              |||||||||||||||||||||||||||||||||||||||||||||
Db        121 CACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACC--------------- 165

Qy        559 ValSerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHis 578

Db        165 ------------------------------------------------------------ 165

Qy        579 CysGlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGly 598

Db        165 ------------------------------------------------------------ 165

Qy        599 SerTrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCys 618

Db        165 ------------------------------------------------------------ 165

Qy        619 GlnLeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuVal 638

Db        165 ------------------------------------------------------------ 165

Qy        639 GlyGluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProVal 658

Db        165 ------------------------------------------------------------ 165

Qy        659 AlaCysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAla 678

Db        165 ------------------------------------------------------------ 165

Qy        679 LeuLysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArg 698

Db        165 ------------------------------------------------------------ 165

Qy        699 ValLeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSer 718
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        166 ---CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGC 222

Qy        719 SerLeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrp 738
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        223 TCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGG 282

Qy        739 AsnGlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThr 758
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        283 AATGGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACT 342

Qy        759 SerAspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSer 778
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        343 AGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGC 402

Qy        779 IleAsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGlu 798
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        403 ATCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAA 462

Qy        799 AlaGlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGln 818
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        463 GCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAG 522

Qy        819 LysIleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAla 838
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        523 AAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCC 582

Qy        839 GlnLysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAla 858
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        583 CAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCC 642

Qy        859 MetIleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAla 878
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        643 ATGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCT 702

Qy        879 AlaAlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGlu 898
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        703 GCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCNGAGGCTGAG 762

Qy        899 Cys 899
              |||
Db        763 TGC 765


Search completed: October  6, 2004, 22:37:02
Job time : 5582 secs
```

GenCore version 5.1.6


OM nucleic - nucleic search, using sw model

Run on: October 6, 2004, 09:39:46 ; Search time 11136 Seconds
(without alignments)
11213.297 Million cell updates/sec

Title: US-09-970-944-1
Perfect score: 2881
Sequence: 1 agctggggctccgggctgag..........ccttccccacaccggggaga 2881

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 2881 | 100.0 | 2881 | 6 | AX527916 | AX527916 Sequence |
| 2 | 2676.4 | 92.9 | 2752 | 6 | AX449572 | AX449572 Sequence |
| 3 | 2615 | 90.8 | 2697 | 6 | AX451652 | AX451652 Sequence |
| 4 | 2402.6 | 83.4 | 3580 | 6 | AX367094 | AX367094 Sequence |
| 5 | 2327.8 | 80.8 | 3992 | 10 | MMU487852 | AJ487852 Mus muscu |
| 6 | 2235.8 | 77.6 | 3014 | 6 | BD057524 | BD057524 Netrin re |
| 7 | 2189.4 | 76.0 | 2697 | 6 | AX268596 | AX268596 Sequence |
| 8 | 2189.4 | 76.0 | 2697 | 10 | RNU87305 | U87305 Rattus norv |
| 9 | 2020.6 | 70.1 | 3844 | 10 | BC058084 | BC058084 Mus muscu |
| 10 | 1691.4 | 58.7 | 2688 | 9 | BC009333 | BC009333 Homo sapi |
| 11 | 1570.4 | 54.5 | 1787 | 6 | BD057525 | BD057525 Netrin re |
| 12 | 1291.2 | 44.8 | 9700 | 6 | AX054976 | AX054976 Sequence |
| 13 | 940.2 | 32.6 | 9299 | 10 | MMU72634 | U72634 Mus musculu |
| 14 | 938.4 | 32.6 | 2962 | 5 | AY187310 | AY187310 Gallus ga |
| 15 | 936.8 | 32.5 | 9328 | 10 | AB118026 | AB118026 Rattus no |
| 16 | 916 | 31.8 | 3646 | 9 | AF055634 | AF055634 Homo sapi |
| 17 | 895 | 31.1 | 3770 | 9 | AY126437 | AY126437 Homo sapi |
| 18 | 889.4 | 30.9 | 4294 | 10 | AK122575 | AK122575 Mus muscu |
| 19 | 889 | 30.9 | 2860 | 6 | AX686445 | AX686445 Sequence |
| 20 | 887.4 | 30.8 | 2860 | 6 | AX686447 | AX686447 Sequence |
| 21 | 875.8 | 30.4 | 2995 | 6 | AX497288 | AX497288 Sequence |
| 22 | 866.4 | 30.1 | 2895 | 6 | AX512281 | AX512281 Sequence |
| 23 | 859.8 | 29.8 | 3672 | 10 | BC057560 | BC057560 Mus muscu |
| 24 | 859.8 | 29.8 | 3672 | 10 | BC048162 | BC048162 Mus muscu |
| 25 | 856.8 | 29.7 | 3933 | 9 | AK128132 | AK128132 Homo sapi |
| 26 | 856.8 | 29.7 | 3935 | 9 | AB096256 | AB096256 Homo sapi |
| 27 | 853.6 | 29.6 | 3884 | 6 | AX464012 | AX464012 Sequence |
| 28 | 853.6 | 29.6 | 3884 | 9 | AY358351 | AY358351 Homo sapi |
| 29 | 821.6 | 28.5 | 3788 | 10 | MMU487853 | AJ487853 Mus muscu |
| 30 | 794.4 | 27.6 | 2838 | 10 | RNU87306 | U87306 Rattus norv |
| 31 | 783.4 | 27.2 | 2831 | 6 | BD057526 | BD057526 Netrin re |
| 32 | 735.8 | 25.5 | 813 | 6 | AX054892 | AX054892 Sequence |
| 33 | 665.6 | 23.1 | 2832 | 5 | AY099459 | AY099459 Xenopus l |
| 34 | 597.6 | 20.7 | 2661 | 6 | AX800717 | AX800717 Sequence |
| 35 | 597.6 | 20.7 | 2868 | 6 | AX800719 | AX800719 Sequence |
| 36 | 573.4 | 19.9 | 604 | 6 | AX451656 | AX451656 Sequence |
| 37 | 560.6 | 19.5 | 3252 | 9 | AY358147 | AY358147 Homo sapi |
| 38 | 557.8 | 19.4 | 7002 | 6 | BD162193 | BD162193 Novel unc |
| 39 | 557.8 | 19.4 | 7002 | 9 | AB055056 | AB055056 Homo sapi |
| 40 | 553 | 19.2 | 3925 | 10 | MMU487854 | AJ487854 Mus muscu |
| 41 | 502.6 | 17.4 | 1438 | 6 | AX054888 | AX054888 Sequence |
| 42 | 467.4 | 16.2 | 2406 | 6 | AX833115 | AX833115 Sequence |
| 43 | 467.4 | 16.2 | 2406 | 9 | AK094595 | AK094595 Homo sapi |
| 44 | 465.8 | 16.2 | 1948 | 6 | AX882354 | AX882354 Sequence |
| 45 | 465.8 | 16.2 | 1948 | 6 | BD159676 | BD159676 Primer fo |

ALIGNMENTS

RESULT 1
AX527916
LOCUS AX527916 2881 bp DNA linear PAT 21-NOV-2002
DEFINITION Sequence 1 from Patent WO0229038.
ACCESSION AX527916
VERSION AX527916.1 GI:25172359
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Herrmann,J.L., Rastelli,L. and Shimkets,R.A.
TITLE Novel proteins and nucleic acids encoding same and antibodies
directed against these proteins

```
 JOURNAL   Patent: WO 0229038-A 1 11-APR-2002;
           Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1..2881
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             100.0%;  Score 2881;  DB 6;  Length 2881;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 2881;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGCTGGGGCTCCGGGCTGAGGCGCTAAAGCCGCCCTCCCGCCCGCGGGGCCCCGCGCCCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGCTGGGGCTCCGGGCTGAGGCGCTAAAGCCGCCCTCCCGCCCGCGGGGCCCCGCGCCCG 60

Qy         61 GCCCGCCCGCCTGCCCGCCCGCGGCCATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCCCGCCCGCCTGCCCGCCCGCGGCCATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCC 120

Qy        121 TGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGG 180

Qy        181 CCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGG 240

Qy        241 ATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGC 300

Qy        301 AGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCA 360

Qy        361 GCACAGACGGGAGCAGTGGTGAGCCGACCATGGAGGTCCGCATTAATGTCTCAAGGCAGC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCACAGACGGGAGCAGTGGTGAGCCGACCATGGAGGTCCGCATTAATGTCTCAAGGCAGC 420

Qy        421 AGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCT 480

Qy        481 CGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCG 540

Qy        541 AGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTC 600

Qy        601 CACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACC 660

Qy        661 CGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCC 720

Qy        721 GCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCA 780

Qy        781 GCGCCTCCGCTGCTGTCATCGTCTACGTGAACGGTGGGTGGTCGACGTGGACCGAGTGGT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GCGCCTCCGCTGCTGTCATCGTCTACGTGAACGGTGGGTGGTCGACGTGGACCGAGTGGT 840

Qy        841 CCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAACC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAACC 900

Qy        901 CGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAATGTCCATGACCGCACCGTCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAATGTCCATGACCGCACCGTCT 960

Qy        961 CCTCTCTGCTTGTCTCTGTGGACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCTCTCTGCTTGTCTCTGTGGACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTG 1020

Qy       1021 GGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAG 1080

Qy       1081 GGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGTACCAGTGACCTCTGTGTAC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGTACCAGTGACCTCTGTGTAC 1140

Qy       1141 ACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCT 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 ACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCT 1200

Qy       1201 GCCTGGTCCTGCTGCTGCTTGTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GCCTGGTCCTGCTGCTGCTTGTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGG 1260

Qy       1261 ACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 ACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGC 1320

Qy       1321 CCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCA 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCA 1380

Qy       1381 CCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCA 1440

Qy       1441 ATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 ATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCA 1500

Qy       1501 CCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGC 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGC 1560

Qy       1561 CCCGAGGCACCAGCAACATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGA 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CCCGAGGCACCAGCAACATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGA 1620

Qy       1621 TCCCTAATACAGGTATCAGCCTCCTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCT 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 TCCCTAATACAGGTATCAGCCTCCTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCT 1680

Qy       1681 ATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTC 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 ATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTC 1740

Qy       1741 AGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAG 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 AGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAG 1800

Qy       1801 TCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCA 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 TCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCA 1860

Qy       1861 AAAAGCAGTCGTGCGAGGGCAGCTGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGC 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 AAAAGCAGTCGTGCGAGGGCAGCTGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGC 1920

Qy       1921 CCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGC 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 CCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGC 1980

Qy       1981 TGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGC 2040
```

```
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 TGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGC 2040

Qy       2041 TGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCC 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 TGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCC 2100

Qy       2101 TGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGAC 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 TGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGAC 2160

Qy       2161 AGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTAT 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 AGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTAT 2220

Qy       2221 CCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGA 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 CCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGA 2280

Qy       2281 TCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGG 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 TCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGG 2340

Qy       2341 AGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGG 2400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 AGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGG 2400

Qy       2401 GCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGC 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 GCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGC 2460

Qy       2461 TGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGA 2520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 TGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGA 2520

Qy       2521 TCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTG 2580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 TCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTG 2580

Qy       2581 CCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCT 2640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 CCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCT 2640

Qy       2641 CCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACG 2700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 CCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACG 2700

Qy       2701 GCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCT 2760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 GCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCT 2760

Qy       2761 TCACAGTGTCGGAGGCTGAGTGCTGAGGCCGGCCAGGCCCGACACCTACACTCTCACCAG 2820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 TCACAGTGTCGGAGGCTGAGTGCTGAGGCCGGCCAGGCCCGACACCTACACTCTCACCAG 2820

Qy       2821 CTTTGGCACCCACCAAGGACAGGCAGAAGCCGGACAGGGGCCCTTCCCCACACCGGGGAG 2880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 CTTTGGCACCCACCAAGGACAGGCAGAAGCCGGACAGGGGCCCTTCCCCACACCGGGGAG 2880

Qy       2881 A 2881
              |
Db       2881 A 2881

RESULT 2
AX449572
LOCUS       AX449572                2752 bp    DNA     linear   PAT 03-JUL-2002
DEFINITION  Sequence 1 from Patent WO0210216.
ACCESSION   AX449572
VERSION     AX449572.1  GI:21698195
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Padigaru,M., Mezes,P., Mishra,V., Burgess,C., Casman,S.,
            Grosse,W.M., Alsobrook,J.P., Lepley,D.M., Gerlach,V.L.,
            Macdougall,J.R. and Smithson,G.
  TITLE     Proteins and nucleic acids encoding same
  JOURNAL   Patent: WO 0210216-A 1 07-FEB-2002;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1..2752
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             92.9%;  Score 2676.4;  DB 6;  Length 2752;
  Best Local Similarity   98.9%;  Pred. No. 0;
  Matches 2728;  Conservative    0;  Mismatches   21;  Indels    9;  Gaps    3;

Qy         42 CCGCGGGGCCCCGCGCCCGGCCCGCCCGCCTGCCCGCCCGCGGCCATGGCCGTCCGGCCC 101
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCGCGGGGCCCCGCGCCCGGCCCGCCCGCCTGCCCGCCCGCGGCCATGGCCGTCCGGCCC 60

Qy        102 GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC 161
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC 120

Qy        162 CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC 221
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC 180

Qy        222 TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC 281
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC 240

Qy        282 AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG 341
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG 300

Qy        342 GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCGACCATGGAGGTCCGC 401
              ||||||||||||||||||||||||||||||||||||||||    |||  |||||||||||||
Db        301 GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCCACCATGGAGGTCCGC 360

Qy        402 ATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAG 461
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAG 420

Qy        462 TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC 521
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC 480

Qy        522 AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC 581
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC 540

Qy        582 ATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG 641
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG 600

Qy        642 AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC 701
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC 660

Qy        702 CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC 761
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC 720

Qy        762 ATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTACGTGAACGGTGGGTGG 821
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTACGTGAACGGTGGGTGG 780

Qy        822 TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG 881
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       781 TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG 840

Qy       882 AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT 941
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       841 AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT 900

Qy       942 GTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGCAGCTGGAGCCCGTGG 1001
             |||||  |    || | || | |    ||   | | || |||||||||||||||||||||
Db       901 GTCCAGAA---AACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGCTGGAGCCCGTGG 957

Qy      1002 AGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC 1061
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       958 AGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC 1017

Qy      1062 CCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT 1121
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1018 CCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT 1077

Qy      1122 ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC 1181
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1078 ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC 1137

Qy      1182 CTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCATCCTCGTTTATTGC 1241
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1138 CTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCATCCTCGTTTATTGC 1197

Qy      1242 CGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC 1301
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1198 CGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC 1257

Qy      1302 CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG 1361
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1258 CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG 1317

Qy      1362 GACCTCAG---CACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC 1418
             ||||||||   |||||||||||||||||||||||||||||||||||||||||||||||||
Db      1318 GACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC 1377

Qy      1419 AGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC 1478
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1378 AGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC 1437

Qy      1479 ACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC 1538
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1438 ACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC 1497

Qy      1539 CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC 1598
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1498 CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC 1557

Qy      1599 TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCCAGAT 1658
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1558 TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCCAGAT 1617

Qy      1659 GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG 1718
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1618 GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG 1677

Qy      1719 AGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT 1778
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1678 AGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT 1737

Qy      1779 GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT 1838
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1738 GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT 1797

Qy      1839 GACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGCAGGATGTG 1898
             |||||||||||||||||||||||||||||||||||||||||||||||||   ||||||||
Db      1798 GACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGG---AGGATGTG 1854

Qy      1899 CTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1855 CTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC 1914

Qy      1959 TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1915 TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG 1974

Qy      2019 GCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1975 GCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAG 2034

Qy      2079 TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG 2138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2035 TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG 2094

Qy      2139 CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC 2198
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2095 CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC 2154

Qy      2199 AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAG 2258
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2155 AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAG 2214

Qy      2259 CTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC 2318
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2215 CTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC 2274

Qy      2319 TTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAG 2378
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2275 TTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAG 2334

Qy      2379 CTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACC 2438
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2335 CTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACC 2394

Qy      2439 AAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG 2498
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2395 AAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG 2454

Qy      2499 GTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTG 2558
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2455 GTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTG 2514

Qy      2559 GACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGAC 2618
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2515 GACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGAC 2574

Qy      2619 AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGG 2678
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2575 AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGG 2634

Qy      2679 GAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTG 2738
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2635 GAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTG 2694

Qy      2739 GGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGAGGCCGGCCAG 2796
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2695 GGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGAGGCCGGCCAG 2752
```

```
RESULT 3
AX451652
LOCUS       AX451652             2697 bp    DNA     linear   PAT 03-JUL-2002
DEFINITION  Sequence 1 from Patent WO0233080.
ACCESSION   AX451652
VERSION     AX451652.1  GI:21698587
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Koehler,R.H.
  TITLE     Regulation of human netrin binding membrane receptor unc5h-1
```

```
JOURNAL    Patent: WO 0233080-A 1 25-APR-2002;
           Bayer Aktiengesellschaft (DE)
FEATURES            Location/Qualifiers
     source         1..2697
                    /organism="Homo sapiens"
                    /mol_type="unassigned DNA"
                    /db_xref="taxon:9606"
ORIGIN

  Query Match             90.8%;  Score 2615;  DB 6;  Length 2697;
  Best Local Similarity   98.7%;  Pred. No. 0;
  Matches 2669;  Conservative    0;  Mismatches   25;  Indels    9;  Gaps    3;

Qy         87 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 146
Db          1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 60

Qy        147 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 206
Db         61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120

Qy        207 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 266
Db        121 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 180

Qy        267 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 326
Db        181 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 240

Qy        327 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCG 386
Db        241 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 300

Qy        387 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 446
Db        301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 360

Qy        447 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 506
Db        361 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 420

Qy        507 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 566
Db        421 TACATCCGCATAGCCTATTTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 480

Qy        567 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 626
Db        481 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 540

Qy        627 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 686
Db        541 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 600

Qy        687 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 746
Db        601 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 660

Qy        747 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 806
Db        661 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 720

Qy        807 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 866
Db        721 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 780

Qy        867 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 926
Db        781 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 840

Qy        927 TGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGC 986
Db        841 TGTGAGGGGCAGAATGTCCAGAA---AACAGCCTGCGCCACCCTGTGCCCAGTGGACGGC 897

Qy        987 AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC 1046
Db        898 AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC 957

Qy       1047 CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG 1106
Db        958 CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG 1017

Qy       1107 GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG 1166
Db       1018 GACACCCGCAACTGTACCAGTGACCTCTGTGTACACACTGCTTCTGGCCCTGAGGACGTG 1077

Qy       1167 GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC 1226
Db       1078 GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC 1137

Qy       1227 ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1286
Db       1138 ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1197

Qy       1287 CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1346
Db       1198 CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1257

Qy       1347 CTCACCATCCAGCCGGACCTCAG---CACCACCACCACCTACCAGGGCAGTCTCTGTCCC 1403
Db       1258 CTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCC 1317

Qy       1404 CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG 1463
Db       1318 CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG 1377

Qy       1464 GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC 1523
Db       1378 GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC 1437

Qy       1524 TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC 1583
Db       1438 TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC 1497

Qy       1584 TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC 1643
Db       1498 TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGAATCAGCCTC 1557

Qy       1644 CTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC 1703
Db       1558 CTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC 1617

Qy       1704 AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT 1763
Db       1618 AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT 1677

Qy       1764 AGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT 1823
Db       1678 AGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT 1737

Qy       1824 GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC 1883
Db       1738 GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC 1797

Qy       1884 TGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG 1943
Db       1798 TGGG---AGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG 1854

Qy       1944 CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA 2003
Db       1855 CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA 1914

Qy       2004 GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC 2063
Db       1915 GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC 1974

Qy       2064 TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC 2123
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1975 TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC 2034

Qy       2124 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGGGTC 2183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2035 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGGGTC 2094

Qy       2184 CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCC 2243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2095 CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCC 2154

Qy       2244 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2155 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2214

Qy       2304 GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2215 GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2274

Qy       2364 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2275 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2334

Qy       2424 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2335 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2394

Qy       2484 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2395 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2454

Qy       2544 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2455 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2514

Qy       2604 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG 2663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2515 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG 2574

Qy       2664 ATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2575 ATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2634

Qy       2724 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2635 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2694

Qy       2784 TGA 2786
              |||
Db       2695 TGA 2697
```

```
RESULT 4
AX367094
LOCUS       AX367094                3580 bp    DNA     linear   PAT 16-FEB-2002
DEFINITION  Sequence 13 from Patent WO0198354.
ACCESSION   AX367094
VERSION     AX367094.1  GI:18855296
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Griffin,J.A., Kallick,D.A., Tribouley,C.M., Yue,H., Nguyen,D.B.,
            Tang,Y.T., Lal,P., Policky,J.L., Azimzai,Y., Lu,D.A., Graul,R.,
            Yao,M.G., Burford,N., Hafalia,A.J., Baughn,M.R., Bandman,O.,
            Patterson,C., Yang,J., Xu,Y., Warren,B.A., Ding,L. and
            Sanjanwala,M.S.
  TITLE     Receptors
  JOURNAL   Patent: WO 0198354-A 13 27-DEC-2001;
            Incyte Genomics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1..3580
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="Incyte ID No: 6052371CB1"
ORIGIN

  Query Match             83.4%;  Score 2402.6;  DB 6;  Length 3580;
  Best Local Similarity   93.4%;  Pred. No. 0;
  Matches 2615;  Conservative    0;  Mismatches    9;  Indels  177;  Gaps    3;

Qy         84 GCCATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGG 143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCCATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGG 60

Qy        144 CTCCGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAAC 203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTCCGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAAC 120

Qy        204 CCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAG 263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAG 180

Qy        264 CCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGG 323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGG 240

Qy        324 GAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAG 383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||   | 
Db        241 GAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTG 300

Qy        384 CCGACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTG 443
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTG 360

Qy        444 GAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAG 503
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAG 420

Qy        504 GCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAG 563
              ||||||||||||||||||    ||||||||||||||||||||||||||||||||||||||
Db        421 GCCTACATCCGCATAGCCTATTTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAG 480

Qy        564 GTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCC 623
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCC 540

Qy        624 GAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATAC 683
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATAC 600

Qy        684 ATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTAC 743
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTAC 660

Qy        744 ACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTC 803
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTC 720

Qy        804 TACGTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGG 863
              |||
Db        721 TAC--------------------------------------------------------- 723

Qy        864 CGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCT 923
                                                                          
Db        724 ------------------------------------------------------------ 723

Qy        924 TTCTGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGAC 983
                                                                     ||||||
Db        724 ------------------------------------------------------GTGGAC 729

Qy        984 GGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGG 1043
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      730 GGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGG 789

Qy     1044 AGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGAC 1103
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      790 AGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGAC 849

Qy     1104 CTGGACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGAC 1163
            |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db      850 CTGGACACCCGCAACTGTACCAGTGACCTCTGTGTACACACTGCTTCTGGCCCTGAGGAC 909

Qy     1164 GTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTC 1223
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      910 GTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTC 969

Qy     1224 CTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCC 1283
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      970 CTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCC 1029

Qy     1284 ATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCAT 1343
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1030 ATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCAT 1089

Qy     1344 CTGCTCACCATCCAGCCGGACCTCAG---CACCACCACCACCTACCAGGGCAGTCTCTGT 1400
            ||||||||||||||||||||||||||   |||||||||||||||||||||||||||||||
Db     1090 CTGCTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGT 1149

Qy     1401 CCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCC 1460
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1150 CCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCC 1209

Qy     1461 CTGGGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTC 1520
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1210 CTGGGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTC 1269

Qy     1521 GTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATG 1580
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1270 GTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATG 1329

Qy     1581 ACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGC 1640
            ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db     1330 ACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGAATCAGC 1389

Qy     1641 CTCCTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTG 1700
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1390 CTCCTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTG 1449

Qy     1701 CACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATC 1760
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1450 CACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATC 1509

Qy     1761 GTTAGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCAC 1820
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1510 GTTAGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCAC 1569

Qy     1821 TGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGC 1880
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1570 TGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGC 1629

Qy     1881 AGCTGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGC 1940
            |||||||   ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1630 AGCTGGG---AGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGC 1686

Qy     1941 CAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTG 2000
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1687 CAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTG 1746

Qy     2001 GGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTG 2060
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1747 GGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTG 1806

Qy     2061 GCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCA 2120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1807 GCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCA 1866

Qy     2121 CTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGG 2180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1867 CTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGG 1926

Qy     2181 GTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGC 2240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1927 GTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGC 1986

Qy     2241 TCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGG 2300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1987 TCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGG 2046

Qy     2301 AATGGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACT 2360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2047 AATGGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACT 2106

Qy     2361 AGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGC 2420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2107 AGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGC 2166

Qy     2421 ATCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAA 2480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2167 ATCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAA 2226

Qy     2481 GCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAG 2540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2227 GCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAG 2286

Qy     2541 AAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCC 2600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2287 AAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCC 2346

Qy     2601 CAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCC 2660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2347 CAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCC 2406

Qy     2661 ATGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCT 2720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2407 ATGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCT 2466

Qy     2721 GCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAG 2780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2467 GCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAG 2526

Qy     2781 TGCTGAGGCCGGCCAGGCCCGACACCTACACTCTCACCAGCTTTGGCACCCACCAAGGAC 2840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2527 TGCTGAGGCCGGCCAGGCCCGACACCTACACTCTCACCAGCTTTGGCACCCACCAAGGAC 2586

Qy     2841 AGGCAGAAGCCGGACAGGGGCCCTTCCCCACACCGGGGAGA 2881
            |||||||||||||||||||||||||||||||||||||||||
Db     2587 AGGCAGAAGCCGGACAGGGGCCCTTCCCCACACCGGGGAGA 2627
```

```
RESULT 5
MMU487852
LOCUS       MMU487852              3992 bp    mRNA    linear   ROD 24-SEP-2002
DEFINITION  Mus musculus mRNA for netrin receptor Unc5h1 (Unc5h1 gene).
ACCESSION   AJ487852
VERSION     AJ487852.1  GI:22035783
KEYWORDS    netrin receptor Unc5h1; Unc5h1 gene.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Engelkamp,D.
  TITLE     Cloning of three mouse Unc5 genes and their expression patterns at
            mid-gestation
  JOURNAL   Mech. Dev. 118 (1-2), 191-197 (2002)
  MEDLINE   22239710
   PUBMED   12351186
REFERENCE   2  (bases 1 to 3992)
```

```
AUTHORS   Engelkamp,D.
TITLE     Direct Submission
JOURNAL   Submitted (15-MAY-2002) Neuroanatomy, Max Planck Institute for
          Brain Research, Deutschordenstrasse 46, Frankfurt 60528, GERMANY
FEATURES             Location/Qualifiers
     source          1..3992
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
     gene            1..3992
                     /gene="Unc5h1"
     CDS             232..2928
                     /gene="Unc5h1"
                     /codon_start=1
                     /product="netrin receptor Unc5h1"
                     /protein_id="CAD32250.1"
                     /db_xref="GI:22035784"
                     /db_xref="GOA:Q8K1S4"
                     /db_xref="SPTREMBL:Q8K1S4"
                     /translation="MAVRPGLWPALLGIVLTAWLRGSGAQQSATVANPVPGANPDLLP
                     HFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTM
                     EVRINVSRQQVEKVFGLEEYWCQCVAWSSSGTTKSQKAYIRIAYLRKNFEQEPLAKEV
                     SLEQGIVLPCRPPEGIPPAEVEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTAN
                     YTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLN
                     GGAFCEGQNVQKTACATLCPVDGSWSPWSKWSACGLDCTHWRSRECSDPAPRNGGEEC
                     RGADLDTRNCTSDLCLHTSSGPEDVALYIGLVAVAVCLILLLLVLVLIYCRKKEGLDS
                     DVADSSILTSGFQPVSIKPSKADNPHLLTIQPDLSTTTTTYQGSLCPRQDGPSPKFQL
                     SNGHLLSPLGSGRHTLHHSSPTSEAEDFVSRLSTQNYFRSLPRGTSNMAYGTFNFLGG
                     RLMIPNTGISLLIPPDAIPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGV
                     LLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGSWEDVLHLGEESPSHLYYCQLEAGAC
                     YVFTEQLGRFALVGEALSVAATKRLRLLLFAPVACTSLEYNIRVYCLHDTHDALKEVV
                     QLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWKSKLLVSYQEIPFYHIWNGT
                     QQYLHCTFTLERVNASTSDLACKVWVWQVEGDGQSFNINFNITKDTRFAEMLALESEG
                     GVPALVGPSAFKIPFLIRQKIITSLDPPCSRGADWRTLAQKLHLDSHLSFFASKPSPT
                     AMILNLWEARHFPNGNLGQLAAAVAGLGQPDAGLFTVSEAEC"
ORIGIN

  Query Match             80.8%;  Score 2327.8;  DB 10;  Length 3992;
  Best Local Similarity   89.3%;  Pred. No. 0;
  Matches 2579;  Conservative    0;  Mismatches  292;  Indels   16;  Gaps    6;

Qy          1 AGCTGGGGCTCCGGGCTGAGGCGCT-AAAGCCGCCCTCCCGCCCGCGGGCCCCGCGCCC 59
Db        149 AGGTGGGGCTCTCGGCTGAGGCGCTGACAGCCTCTCTCCCGCCCGCGGGCCCCTAGTCC 208

Qy         60 GGCCCGCCCGCCTGCCCGCCCGCGGCCATGGCCGTCCGGCCCGGCTGTGGCCAGCGCTC 119
Db        209 AGCCCGCTCGT----CCGCCCGCGGCCATGGCCGTCCGGCCCGGCTGTGGCCAGCGCTC 264

Qy        120 CTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTG 179
Db        265 CTGGGCATAGTCCTCACTGCCTGGCTTCGTGGTTCGGGTGCCCAGCAGAGTGCCACAGTG 324

Qy        180 GCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAG 239
Db        325 GCCAACCCAGTGCCTGGTGCCAACCCAGACCTGCTGCCCCACTTCCTGGTAGAGCCGGAG 384

Qy        240 GATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACG 299
Db        385 GACGTGTACATTGTCAAGAACAAGCCCGTGCTGCTGGTGTGCAAGGCTGTGCCCGCCACC 444

Qy        300 CAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGC 359
Db        445 CAGATCTTCTTCAAGTGCAACGGGGAATGGGTTCGCCAGGTCGATCACGTCATCGAACGC 504

Qy        360 AGCACAGACGGGAGCAGTGGTGAGCCGACCATGGAGGTCCGCATTAATGTCTCAAGGCAG 419
Db        505 AGCACTGACGGCAGCAGCGGATTGCCAACCATGGAGGTCCGGATCAACGTATCAAGGCAG 564

Qy        420 CAGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCC 479
Db        565 CAGGTCGAGAAAGTGTTTGGGCTGGAGGAGTACTGGTGCCAGTGTGTGGCATGGAGCTCC 624

Qy        480 TCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCCAGATTGCGCAAGAACTTC 539
Db        625 TCAGGAACCACCAAAAGCCAGAAGGCCTACATCCGGATTGCCTATTTGCGCAAGAACTTT 684

Qy        540 GAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGT 599
Db        685 GAGCAGGAGCCGCTGGCCAAGGAAGTGTCACTGGAGCAAGGCATTGTGCTACCTTGTCGC 744

Qy        600 CCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGAC 659
Db        745 CCCCCGGAAGGAATCCCCCCAGCTGAGGTGGAGTGGCTCCGAAATGAGGACCTCGTGGAC 804

Qy        660 CCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCC 719
Db        805 CCCTCCCTCGACCCCAATGTGTACATCACACGGGAGCACAGCCTAGTCGTGCGGCAGGCC 864

Qy        720 CGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGC 779
Db        865 CGCCTGGCCGACACTGCCAACTACACCTGCGTGGCCAAGAACATCGTGGCCCGTCGCCGA 924

Qy        780 AGCGCCTCCGCTGCTGTCATCGTCTACGTGAACGGTGGGTGGTCGACGTGGACCGAGTGG 839
Db        925 AGCGCCTCTGCGGCCGTCATTGTTTATGTGAACGGTGGGTGGTCGACGTGGACCGAGTGG 984

Qy        840 TCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAAC 899
Db        985 TCCGTCTGCAGTGCCAGCTGTGGGCGTGGCTGGCAGAAACGGAGCCGGAGCTGCACCAAC 1044

Qy        900 CCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAATGTCCATGACCGCACCGTC 959
Db       1045 CCGGCACCTCTCAACGGGGGCGCCTTCTGTGAGGGGCAGAATGTCCAGAAAACAGCCTGC 1104

Qy        960 TCCTCTCTGCTTGTCTCTGTGGACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGT 1019
Db       1105 GCCACTCTG---TGCCCAGTGGATGGGAGCTGGAGCCCATGGAGTAAGTGGTCAGCCTGC 1161

Qy       1020 GGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGA 1079
Db       1162 GGGCTTGACTGCACCCACTGGCGGAGCCGGGAGTGCTCCGACCCAGCGCCCCGCAACGGA 1221

Qy       1080 GGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGTACCAGTGACCTCTGTGTA 1139
Db       1222 GGTGAGGAGTGCCGGGGTGCTGACCTGGACACCCGCAACTGTACCAGTGACCTCTGCCTG 1281

Qy       1140 CACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTC 1199
Db       1282 CACACCTCTTCCGGCCCCGAGGACGTGGCTCTCTACATCGGCCTCGTCGCCGTGGCCGTG 1341

Qy       1200 TGCCTGGTCCTGCTGCTGCTTGTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTG 1259
Db       1342 TGCCTCATCTTGCTGCTGCTGGTCCTCGTCCTCATCTACTGCCGCAAGAAGGAAGGACTG 1401

Qy       1260 GACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAG 1319
Db       1402 GACTCAGACGTGGCTGACTCATCCATCCTTACCTCAGGCTTCCAGCCTGTCAGCATCAAG 1461

Qy       1320 CCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCGGACCTCAGCACCACC--- 1376
Db       1462 CCCAGCAAAGCAGACAATCCCCATCTGCTCACCATCCAACCGGACCTCAGCACCACCACG 1521

Qy       1377 ACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTC 1436
Db       1522 ACCACCTACCAGGGCAGCCTGTGTCCCCGGCAGGATGGACCCAGCCCCAAGTTCCAGCTC 1581

Qy       1437 ACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCACACACTGCACCACAGCTCT 1496
Db       1582 TCTAATGGTCACCTGCTCAGCCCACTGGGCAGTGGCCGCCATACGCTGCACCACAGCTCC 1641

Qy       1497 CCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCC 1556
Db       1642 CCCACCTCTGAGGCTGAGGACTTCGTCTCCCGCCTCTCCACCCAAAACTACTTTCGTTCT 1701

Qy       1557 CTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTG 1616
```

```
Db   1702 CTGCCCCGCGGTACCAGCAACATGGCCTATGGACCTTCAACTTCCTCGGGGGCCGGCTG 1761

Qy   1617 ATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCCAGATGCCATACCCCGAGGGAAG 1676
Db   1762 ATGATCCCTAACACAGGAATCAGCCTCCTCATACCCCCGGACGCCATCCCCCGAGGAAAG 1821

Qy   1677 ATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGC 1736
Db   1822 ATCTACGAGATCTACCTCACTCTGCACAAGCCAGAAGACGTGAGGTTGCCCCTAGCTGGC 1881

Qy   1737 TGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCTGGCGTCCTGCTCACCCGG 1796
Db   1882 TGTCAGACCCTGCTGAGTCCTATCGTTAGCTGTGGGCCCCCAGGAGTCCTGCTCACCCGG 1941

Qy   1797 CCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGC 1856
Db   1942 CCAGTCATCCTTGCCATGGACCACTGCGGGGAGCCCAGTCCCGACAGCTGGAGCCTGCGC 2001

Qy   1857 CTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGCAGGATGTGCTGCACCTGGGCGAGGAG 1916
Db   2002 CTCAAAAAGCAGTCCTGTGAGGGCAGCTGGG---AGGACGTGCTGCACCTTGGTGAGGAG 2058

Qy   1917 GCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAG 1976
Db   2059 TCGCCCTCTCATCTCTACTACTGCCAGCTGGAGGCCGGGGCCTGCTATGTCTTCACCGAG 2118

Qy   1977 CAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTC 2036
Db   2119 CAGCTAGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCACCAAGCGCCTC 2178

Qy   2037 AAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTAC 2096
Db   2179 AGGCTCCTTCTGTTTGCCCCTGTGGCCTGTACGTCCCTCGAGTACAACATCCGAGTGTAC 2238

Qy   2097 TGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGG 2156
Db   2239 TGCCTGCACGACACCCACGATGCTCTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGT 2298

Qy   2157 GGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGC 2216
Db   2299 GGACAGCTGATCCAGGAGCCCCGTGTCCTGCACTTCAAAGACAGTTACCACAACCTACGT 2358

Qy   2217 CTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAG 2276
Db   2359 CTGTCCATCCACGACGTGCCCAGCTCCCTGTGGAAGAGCAAGCTCCTTGTCAGCTACCAG 2418

Qy   2277 GAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTTGCACTGCACCTTCACC 2336
Db   2419 GAGATCCCTTTTTACCACATCTGGAATGGCACTCAGCAGTATCTGCACTGCACCTTCACC 2478

Qy   2337 CTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTG 2396
Db   2479 CTGGAGCGCGTCAATGCCAGCACCAGCGACCTGGCCTGCAAGGTGTGGGTGTGGCAGGTG 2538

Qy   2397 GAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACCAAGGACACAAGGTTTGCT 2456
Db   2539 GAGGGAGATGGACAGAGCTTCAACATCAACTTTAACATCACTAAGGACACGAGGTTTGCT 2598

Qy   2457 GAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTC 2516
Db   2599 GAAATGCTGGCTCTGGAGAGTGAAGGGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTC 2658

Qy   2517 AAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGG 2576
Db   2659 AAGATCCCCTTCCTCATTCGGCAAAAGATCATTACCAGCCTGGACCCACCCTGCAGCCGG 2718

Qy   2577 GGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTT 2636
Db   2719 GGCGCCGACTGGCGAACTCTAGCCCAGAAACTTCACCTGGACAGCCATCTTAGCTTCTTT 2778

Qy   2637 GCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCC 2696
Db   2779 GCCTCCAAGCCCAGCCCTACAGCCATGATCCTCAACCTATGGGAGGCGCGGCACTTCCCC 2838

Qy   2697 AACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGC 2756
Db   2839 AACGGCAACCTCGGCCAGCTGGCCGCAGCTGTGGCCGGACTGGGCCAGCCAGATGCTGGC 2898

Qy   2757 CTCTTCACAGTGTCGGAGGCTGAGTGCTGAGGCCGGCCAGGCCCG--ACACCTACACTCT 2814
Db   2899 CTCTTCACCGTGTCAGAGGCCGAGTGCTGAGACCAGCCAGGCCGGTCACGCCTACACTCT 2958

Qy   2815 CACCAGCTTTGGCACCCACCAAGGACAGGCAGAAGCCGGACAGGGGCCCTTCCCCACACC 2874
Db   2959 CACCAGCTTTGGCACCTGCCAGGGACAGGCAAAAGCCAGACAGGGGCCCTACCCCCACAC 3018

Qy   2875 GGGGAGA 2881
Db   3019 CCGGGGA 3025
```

RESULT 6
BD057524

```
LOCUS       BD057524                3014 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Netrin receptors.
ACCESSION   BD057524
VERSION     BD057524.1  GI:22603130
KEYWORDS    JP 2001505062-A/1.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 3014)
  AUTHORS   Lavigne,M.T., Leonardo,D.E., Hinck,L., Masu,M. and Masu,K.K.
  TITLE     Netrin receptors
  JOURNAL   Patent: JP 2001505062-A 1 17-APR-2001;
            THE REGENTS OF THE UNIV OF CALIFORNIA
COMMENT     PN   JP 2001505062-A/1
            PD   17-APR-2001
            PF   19-FEB-1998 JP 1998536840
            PR   19-FEB-1997 US     08/808982
            PI   MARC TESSIER LAVIGNE,DAVID E LEONARDO,LINDSAY HINCK,MASAYUKI
            PI   MASU,
            PI   KAZUKO KEINO MASU
            PC   C07K1/00,C07K14/00,C07K17/00,C07H21/02,C07H21/04,G01N33/53 CC
             Strandedness: Double;
            CC   Topology: Linear;
            FH   Key             Location/Qualifiers.
FEATURES             Location/Qualifiers
     source          1..3014
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match             77.6%;  Score 2235.8;  DB 6;  Length 3014;
  Best Local Similarity   88.5%;  Pred. No. 0;
  Matches 2474;  Conservative    0;  Mismatches  312;  Indels   11;  Gaps    4;

Qy         87 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 146
Db          1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGTGCTCCTGGGCATAGTCCTCGCCGCCTGGCTT 60

Qy        147 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 206
Db         61 CGTGGTTCGGGTGCCCAGCAGAGTGCCACGGTGGCCAATCCAGTGCCCGGTGCCAACCCC 120

Qy        207 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 266
Db        121 GACCTGCTGCCCCACTTCCTGGTAGAGCCTGAGGACGTGTACATTGTCAAGAACAAGCCG 180

Qy        267 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 326
Db        181 GTGTTGTTGGTGTGCAAGGCTGTGCCTGCCACCCAGATCTTCTTCAAGTGCAATGGGGAA 240

Qy        327 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCG 386
```

```
Db    241 TGGGTCCGCCAGGTCGATCACGTAATTGAACGCAGCACCGACAGCAGCAGCGGATTGCCA 300

Qy    387 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 446
Db    301 ACCATGGAGGTCCGTATCAACGTATCGAGGCAGCAGGTAGAGAAAGTGTTTGGGCTGGAG 360

Qy    447 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 506
Db    361 GAATACTGGTGCCAGTGTGTGGCATGGAGCTCCTCGGGTACCACCAAAAGTCAGAAGGCC 420

Qy    507 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 566
Db    421 TACATCCGGATTGCCTATTTGCGCAAGAACTTTGAGCAGGAGCCACTGGCCAAGGAAGTG 480

Qy    567 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 626
Db    481 TCACTGGAGCAAGGCATTGTACTACCTTGTCGCCCCCCAGAAGGAATCCCCCCAGCTGAG 540

Qy    627 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 686
Db    541 GTGGAGTGGCTTCGAAATGAGGACCTCGTGGACCCCTCCCTCGATCCCAATGTGTACATC 600

Qy    687 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 746
Db    601 ACGCGGGAGCACAGCCTAGTCGTGCGTCAGGCCCGCCTGGCCGACACGGCCAACTACACC 660

Qy    747 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 806
Db    661 TGTGTGGCCAAGAACATCGTAGCCCGTCGCCGAAGCACCTCTGCAGCGGTCATTGTTTAT 720

Qy    807 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 866
Db    721 GTGAACGGTGGGTGGTCGACGTGGACTGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGT 780

Qy    867 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 926
Db    781 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCACCTCTCAACGGGGGCGCCTTC 840

Qy    927 TGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGC 986
Db    841 TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACTCTG---TGCCCAGTGGATGGG 897

Qy    987 AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC 1046
Db    898 AGCTGGAGTTCGTGGAGTAAGTGGTCAGCCTGTGGGCTTGACTGCACCCACTGGCGGAGC 957

Qy   1047 CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG 1106
Db    958 CGCGAGTGCTCTGACCCAGCACCCCGCAATGGAGGTGAGGAGTGTCGGGGTGCTGACCTG 1017

Qy   1107 GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG 1166
Db   1018 GACACCCGCAACTGTACCAGTGACCTCTGCCTGCACACCGCTTCTTGCCCCGAGGACGTG 1077

Qy   1167 GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC 1226
Db   1078 GCTCTCTACATCGGCCTTGTCGCTGTGGCTGTGTGCCTCTTCTTGCTGTTGCTGGCCCTT 1137

Qy   1227 ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1286
Db   1138 GGACTCATTTACTGTCGCAAGAAGGAAGGGCTGGACTCCGATGTGGCCGACTCGTCCATC 1197

Qy   1287 CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1346
Db   1198 CTCACCTCGGGCTTCCAGCCTGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCACCTG 1257

Qy   1347 CTCACCATCCAGCCGGACCTCAGCACCACC---ACCACCTACCAGGGCAGTCTCTGTCCC 1403
Db   1258 CTCACCATCCAGCCAGACCTCAGCACCACCACTACCACCTACCAGGGCAGTCTATGTTCG 1317

Qy   1404 CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG 1463
Db   1318 AGGCAGGATGGACCCAGCCCCAAGTTCCAGCTCTCTAATGGTCACCTGCTCAGCCCACTG 1377

Qy   1464 GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC 1523
Db   1378 GGGAGTGGCCGCCATACGTTGCACCACAGCTCACCCACCTCTGAGGCTGAGGACTTCGTC 1437

Qy   1524 TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC 1583
Db   1438 TCCCGCCTCTCCACCCAAAACTACTTTCGTTCCCTGCCCCGCGGCACCAGCAACATGGCC 1497

Qy   1584 TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC 1643
Db   1498 TACGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACGGGGATCAGCCTC 1557

Qy   1644 CTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC 1703
Db   1558 CTCATACCCCCGGATGCCATCCCCCGAGGAAAGATCTACGAGATCTACCTCACACTGCAC 1617

Qy   1704 AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT 1763
Db   1618 AAGCCAGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCAGTCGTT 1677

Qy   1764 AGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT 1823
Db   1678 AGCTGTGGGCCCCCAGGAGTCCTGCTCACCCGGCCAGTCATCCTTGCAATGGACCACTGT 1737

Qy   1824 GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC 1883
Db   1738 GGAGAGCCCAGCCCTGACAGCTGGAGTCTGCGCCTCAAAAAGCAGTCCTGCGAGGGCAGT 1797

Qy   1884 TGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG 1943
Db   1798 TGGG---AGGATGTGCTGCACCTTGGTGAGGAGTCACCTTCCCACCTCTACTACTGCCAG 1854

Qy   1944 CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA 2003
Db   1855 CTGGAGGCCGGGGCCTGCTATGTCTTCACGGAGCAGCTGGGCCGCTTTGCCCTGGTAGGA 1914

Qy   2004 GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC 2063
Db   1915 GAGGCCCTCAGCGTGGCTGCCACCAAGCGCCTCAGGCTCCTTCTGTTTGCTCCCGTGGCC 1974

Qy   2064 TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC 2123
Db   1975 TGTACGTCCCTTGAGTACAACATCCGAGTGTACTGCCTACACGACACCCACGACGCTCTC 2034

Qy   2124 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTC 2183
Db   2035 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTAGGTGGACAGCTGATCCAGGAGCCTCGCGTC 2094

Qy   2184 CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCC 2243
Db   2095 CTGCACTTCAAAGACAGTTACCACAACCTACGTCTCTCCATCCACGACGTGCCCAGCTCC 2154

Qy   2244 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2303
Db   2155 CTGTGGAAGAGCAAGCTACTTGTCAGCTACCAGGAGATCCCTTTTTACCACATCTGGAAC 2214

Qy   2304 GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2363
Db   2215 GGCACCCAGCAGTATCTGCACTGCACCTTCACCCTGGAGCGCATCAACGCCAGCACCAGC 2274

Qy   2364 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2423
Db   2275 GACCTGGCCTGCAAGGTGTGGGTGTGGCAGGTGGAGGGAGATGGGCAGAGCTTCAACATC 2334

Qy   2424 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2483
Db   2335 AACTTCAACATCACTAAGGACACAAGGTTTGCTGAATTGTTGGCTCTGGAGAGTGAAGGG 2394

Qy   2484 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2543
Db   2395 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAAAAG 2454
```

```
Qy     2544 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2603
Db     2455 ATCATCGCCAGTCTGGACCCACCCTGCAGCCGGGGCGCCGACTGGAGAACTCTAGCCCAG 2514

Qy     2604 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG 2663
Db     2515 AAACTTCACCTGGACAGCCATCTTAGCTTCTTTGCCTCCAAGCCCAGCCCTACAGCCATG 2574

Qy     2664 ATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2723
Db     2575 ATCCTCAACCTATGGGAGGCACGGCACTTCCCCAACGGCAACCTCGGCCAGCTGGCAGCA 2634

Qy     2724 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2783
Db     2635 GCTGTGGCCGGACTGGGCCAACCAGATGCTGGCCTCTTCACGGTGTCGGAGGCCGAGTGT 2694

Qy     2784 TGAGGCCGGCCAGGCCCG--ACACCTACACTCTCACCAGCTTTGGCACCCACCAAGGACA 2841
Db     2695 TGAGACCAGCCAGGCCGGTAATGCCTACATTCTCACCAGCTTTGACACCTGCCAGGGACA 2754

Qy     2842 GGCAGAAGCCGGACAGGGGCCCTTCCCCACACCGGGG 2878
Db     2755 GGCAAAACCAGACAGGGGCCCTTCCCCCACACCCGGG 2791
```

```
RESULT 7
AX268596
LOCUS       AX268596                2697 bp    DNA     linear   PAT 29-OCT-2001
DEFINITION  Sequence 15 from Patent WO0175440.
ACCESSION   AX268596
VERSION     AX268596.1  GI:16541710
KEYWORDS    .
SOURCE      Rattus sp.
  ORGANISM  Rattus sp.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1
  AUTHORS   Cochran,S.W., Paterson,G.Y., Ohashi,Y.W., Morris,B.Y. and
            Pratt,J.Y.
  TITLE     Schizophrenia related genes
  JOURNAL   Patent: WO 0175440-A 15 11-OCT-2001;
            WELFIDE CORPORATION (JP)
FEATURES             Location/Qualifiers
     source          1..2697
                     /organism="Rattus sp."
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10118"
ORIGIN

  Query Match            76.0%;  Score 2189.4;  DB 6;  Length 2697;
  Best Local Similarity  88.9%;  Pred. No. 0;
  Matches 2403;  Conservative    0;  Mismatches  291;  Indels    9;  Gaps    3;

Qy       87 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 146
Db        1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGTGCTCCTGGGCATAGTCCTCGCCGCCTGGCTT 60

Qy      147 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 206
Db       61 CGTGGTTCGGGTGCCCAGCAGAGTGCCACGGTGGCCAATCCAGTGCCCGGTGCCAACCCC 120

Qy      207 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 266
Db      121 GACCTGCTGCCCCACTTCCTGGTAGAGCCTGAGGACGTGTACATTGTCAAGAACAAGCCG 180

Qy      267 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 326
Db      181 GTGTTGTTGGTGTGCAAGGCTGTGCCTGCCACCCAGATCTTCTTCAAGTGCAATGGGGAA 240

Qy      327 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCG 386
Db      241 TGGGTCCGCCAGGTCGATCACGTAATTGAACGCAGCACCGACAGCAGCAGCGGATTGCCA 300

Qy      387 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 446
Db      301 ACCATGGAGGTCCGTATCAACGTATCGAGGCAGCAGGTAGAGAAAGTGTTTGGGCTGGAG 360

Qy      447 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 506
Db      361 GAATACTGGTGCCAGTGTGTGGCATGGAGCTCCTCGGGTACCACCAAAAGTCAGAAGGCC 420

Qy      507 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 566
Db      421 TACATCCGGATTGCCTATTTGCGCAAGAACTTTGAGCAGGAGCCACTGGCCAAGGAAGTG 480

Qy      567 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 626
Db      481 TCACTGGAGCAAGGCATTGTACTACCTTGTCGCCCCCCAGAAGGAATCCCCCCAGCTGAG 540

Qy      627 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 686
Db      541 GTGGAGTGGCTTCGAAATGAGGACCTCGTGGACCCCTCCCTCGATCCCAATGTGTACATC 600

Qy      687 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 746
Db      601 ACGCGGGAGCACAGCCTAGTCGTGCGTCAGGCCCGCCTGGCCGACACGGCCAACTACACC 660

Qy      747 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 806
Db      661 TGTGTGGCCAAGAACATCGTAGCCCGTCGCCGAAGCACCTCTGCAGCGGTCATTGTTTAT 720

Qy      807 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 866
Db      721 GTGAACGGTGGGTGGTCGACGTGGACTGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGT 780

Qy      867 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 926
Db      781 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCACCTCTCAACGGGGGCGCCTTC 840

Qy      927 TGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGC 986
Db      841 TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACTCTG---TGCCCAGTGGATGGG 897

Qy      987 AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC 1046
Db      898 AGCTGGAGTTCGTGGAGTAAGTGGTCAGCCTGTGGGCTTGACTGCACCCACTGGCGGAGC 957

Qy     1047 CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG 1106
Db      958 CGCGAGTGCTCTGACCCAGCACCCCGCAATGGAGGTGAGGAGTGTCGGGGTGCTGACCTG 1017

Qy     1107 GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG 1166
Db     1018 GACACCCGCAACTGTACCAGTGACCTCTGCCTGCACACCGCTTCTTGCCCCGAGGACGTG 1077

Qy     1167 GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC 1226
Db     1078 GCTCTCTACATCGGCCTTGTCGCTGTGGCTGTGTGCCTCTTCTTGCTGTTGCTGGCCCTT 1137

Qy     1227 ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1286
Db     1138 GGACTCATTTACTGTCGCAAGAAGGAAGGGCTGGACTCCGATGTGGCCGACTCGTCCATC 1197

Qy     1287 CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1346
Db     1198 CTCACCTCGGGCTTCCAGCCTGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCACCTG 1257

Qy     1347 CTCACCATCCAGCCGGACCTCAGCACCACC---ACCACCTACCAGGGCAGTCTCTGTCCC 1403
Db     1258 CTCACCATCCAGCCAGACCTCAGCACCACCACTACCACCTACCAGGGCAGTCTATGTTCG 1317

Qy     1404 CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG 1463
Db     1318 AGGCAGGATGGACCCAGCCCCAAGTTCCAGCTCTCTAATGGTCACCTGCTCAGCCCACTG 1377
```

```
Qy     1464 GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC 1523
Db     1378 GGGAGTGGCCGCCATACGTTGCACCACAGCTCACCCACCTCTGAGGCTGAGGACTTCGTC 1437

Qy     1524 TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC 1583
Db     1438 TCCCGCCTCTCCACCCAAAACTACTTTCGTTCCCTGCCCCGCGGCACCAGCAACATGGCC 1497

Qy     1584 TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC 1643
Db     1498 TACGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACGGGGATCAGCCTC 1557

Qy     1644 CTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC 1703
Db     1558 CTCATACCCCCGGATGCCATCCCCCGAGGAAAGATCTACGAGATCTACCTCACACTGCAC 1617

Qy     1704 AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT 1763
Db     1618 AAGCCAGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCAGTCGTT 1677

Qy     1764 AGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT 1823
Db     1678 AGCTGTGGGCCCCCAGGAGTCCTGCTCACCCGGCCAGTCATCCTTGCAATGGACCACTGT 1737

Qy     1824 GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC 1883
Db     1738 GGAGAGCCCAGCCCTGACAGCTGGAGTCTGCGCCTCAAAAAGCAGTCCTGCGAGGGCAGT 1797

Qy     1884 TGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG 1943
Db     1798 TGGG---AGGATGTGCTGCACCTTGGTGAGGAGTCACCTTCCCACCTCTACTACTGCCAG 1854

Qy     1944 CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA 2003
Db     1855 CTGGAGGCCGGGGCCTGCTATGTCTTCACGGAGCAGCTGGGCCGCTTTGCCCTGGTAGGA 1914

Qy     2004 GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC 2063
Db     1915 GAGGCCCTCAGCGTGGCTGCCACCAAGCGCCTCAGGCTCCTTCTGTTTGCTCCCGTGGCC 1974

Qy     2064 TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC 2123
Db     1975 TGTACGTCCCTTGAGTACAACATCCGAGTGTACTGCCTACACGACACCCACGACGCTCTC 2034

Qy     2124 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTC 2183
Db     2035 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTAGGTGGACAGCTGATCCAGGAGCCTCGCGTC 2094

Qy     2184 CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCC 2243
Db     2095 CTGCACTTCAAAGACAGTTACCACAACCTACGTCTCTCCATCCACGACGTGCCCAGCTCC 2154

Qy     2244 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2303
Db     2155 CTGTGGAAGAGCAAGCTACTTGTCAGCTACCAGGAGATCCCTTTTTACCACATCTGGAAC 2214

Qy     2304 GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2363
Db     2215 GGCACCCAGCAGTATCTGCACTGCACCTTCACCCTGGAGCGCATCAACGCCAGCACCAGC 2274

Qy     2364 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2423
Db     2275 GACCTGGCCTGCAAGGTGTGGGTGTGGCAGGTGGAGGGAGATGGGCAGAGCTTCAACATC 2334

Qy     2424 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2483
Db     2335 AACTTCAACATCACTAAGGACACAAGGTTTGCTGAATTGTTGGCTCTGGAGAGTGAAGGG 2394

Qy     2484 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2543
Db     2395 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAAAAG 2454

Qy     2544 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2603
Db     2455 ATCATCGCCAGTCTGGACCCACCCTGCAGCCGGGGCGCCGACTGGAGAACTCTAGCCCAG 2514

Qy     2604 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG 2663
Db     2515 AAACTTCACCTGGACAGCCATCTTAGCTTCTTTGCCTCCAAGCCCAGCCCTACAGCCATG 2574

Qy     2664 ATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2723
Db     2575 ATCCTCAACCTATGGGAGGCACGGCACTTCCCCAACGGCAACCTCGGCCAGCTGGCAGCA 2634

Qy     2724 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2783
Db     2635 GCTGTGGCCGGACTGGGCCAACCAGATGCTGGCCTCTTCACGGTGTCGGAGGCCGAGTGT 2694

Qy     2784 TGA 2786
Db     2695 TGA 2697
```

```
RESULT 8
RNU87305
LOCUS       RNU87305                2697 bp    mRNA    linear   ROD 15-MAY-1997
DEFINITION  Rattus norvegicus transmembrane receptor Unc5H1 mRNA, complete cds.
ACCESSION   U87305
VERSION     U87305.1  GI:2055391
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 2697)
  AUTHORS   Leonardo,E.D., Hinck,L., Masu,M., Keino-Masu,K., Ackerman,S.L. and
            Tessier-Lavigne,M.
  TITLE     Vertebrate homologues of C. elegans UNC-5 are candidate netrin
            receptors
  JOURNAL   Nature 386 (6627), 833-838 (1997)
  MEDLINE   97271897
   PUBMED   9126742
REFERENCE   2  (bases 1 to 2697)
  AUTHORS   Leonardo,E.D., Hinck,L., Masu,M., Keino-Masu,K. and
            Tessier-Lavigne,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-JAN-1997) Anatomy, UCSF, 513 Parnassus, San
            Francisco, CA 94143-0452, USA
FEATURES             Location/Qualifiers
     source          1..2697
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10116"
                     /tissue_type="brain and ventral spinal cord"
                     /dev_stage="18 day embyo and 13 day embryo"
     CDS             1..2697
                     /codon_start=1
                     /product="transmembrane receptor UNC5H1"
                     /protein_id="AAB57678.1"
                     /db_xref="GI:2055392"
                     /translation="MAVRPGLWPVLLGIVLAAWLRGSGAQQSATVANPVPGANPDLLP
                     HFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDSSSGLPTM
                     EVRINVSRQQVEKVFGLEEYWCQCVAWSSSGTTKSQKAYIRIAYLRKNFEQEPLAKEV
                     SLEQGIVLPCRPPEGIPPAEVEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTAN
                     YTCVAKNIVARRRSTSAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLN
                     GGAFCEGQNVQKTACATLCPVDGSWSSWSKWSACGLDCTHWRSRECSDPAPRNGGEEC
                     RGADLDTRNCTSDLCLHTASCPEDVALYIGLVAVAVCLFLLLLALGLIYCRKKEGLDS
                     DVADSSILTSGFQPVSIKPSKADNPHLLTIQPDLSTTTTTYQGSLCSRQDGPSPKFQL
                     SNGHLLSPLGSGRHTLHHSSPTSEAEDFVSRLSTQNYFRSLPRGTSNMAYGTFNFLGG
                     RLMIPNTGISLLIPPDAIPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPVVSCGPPGV
                     LLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGSWEDVLHLGEESPSHLYYCQLEAGAC
                     YVFTEQLGRFALVGEALSVAATKRLRLLLFAPVACTSLEYNIRVYCLHDTHDALKEVV
                     QLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWKSKLLVSYQEIPFYHIWNGT
                     QQYLHCTFTLERINASTSDLACKVWVWQVEGDGQSFNINFNITKDTRFAELLALESEG
```

```
          GVPALVGPSAFKIPFLIRQKIIASLDPPCSRGADWRTLAQKLHLDSHLSFFASKPSPT
          AMILNLWEARHFPNGNLGQLAAAVAGLGQPDAGLFTVSEAEC"
ORIGIN

  Query Match           76.0%;  Score 2189.4;  DB 10;  Length 2697;
  Best Local Similarity 88.9%;  Pred. No. 0;
  Matches 2403;  Conservative    0;  Mismatches  291;  Indels    9;  Gaps    3;

Qy     87 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 146
Db      1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGTGCTCCTGGGCATAGTCCTCGCCGCCTGGCTT 60

Qy    147 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 206
Db     61 CGTGGTTCGGGTGCCCAGCAGAGTGCCACGGTGGCCAATCCAGTGCCCGGTGCCAACCCC 120

Qy    207 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 266
Db    121 GACCTGCTGCCCCACTTCCTGGTAGAGCCTGAGGACGTGTACATTGTCAAGAACAAGCCG 180

Qy    267 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 326
Db    181 GTGTTGTTGGTGTGCAAGGCTGTGCCTGCCACCCAGATCTTCTTCAAGTGCAATGGGGAA 240

Qy    327 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCG 386
Db    241 TGGGTCCGCCAGGTCGATCACGTAATTGAACGCAGCACCGACAGCAGCAGCGGATTGCCA 300

Qy    387 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 446
Db    301 ACCATGGAGGTCCGTATCAACGTATCGAGGCAGCAGGTAGAGAAAGTGTTTGGGCTGGAG 360

Qy    447 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 506
Db    361 GAATACTGGTGCCAGTGTGTGGCATGGAGCTCCTCGGGTACCACCAAAAGTCAGAAGGCC 420

Qy    507 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 566
Db    421 TACATCCGGATTGCCTATTTGCGCAAGAACTTTGAGCAGGAGCCACTGGCCAAGGAAGTG 480

Qy    567 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 626
Db    481 TCACTGGAGCAAGGCATTGTACTACCTTGTCGCCCCCCAGAAGGAATCCCCCCAGCTGAG 540

Qy    627 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 686
Db    541 GTGGAGTGGCTTCGAAATGAGGACCTCGTGGACCCCTCCCTCGATCCCAATGTGTACATC 600

Qy    687 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 746
Db    601 ACGCGGGAGCACAGCCTAGTCGTGCGTCAGGCCCGCCTGGCCGACACGGCCAACTACACC 660

Qy    747 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 806
Db    661 TGTGTGGCCAAGAACATCGTAGCCCGTCGCCGAAGCACCTCTGCAGCGGTCATTGTTTAT 720

Qy    807 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 866
Db    721 GTGAACGGTGGGTGGTCGACGTGGACTGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGT 780

Qy    867 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 926
Db    781 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCACCTCTCAACGGGGGCGCCTTC 840

Qy    927 TGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGC 986
Db    841 TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACTCTG---TGCCCAGTGGATGGG 897

Qy    987 AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC 1046
Db    898 AGCTGGAGTTCGTGGAGTAAGTGGTCAGCCTGTGGGCTTGACTGCACCCACTGGCGGAGC 957

Qy   1047 CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG 1106
Db    958 CGCGAGTGCTCTGACCCAGCACCCCGCAATGGAGGTGAGGAGTGTCGGGGTGCTGACCTG 1017

Qy   1107 GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG 1166
Db   1018 GACACCCGCAACTGTACCAGTGACCTCTGCCTGCACACCGCTTCTTGCCCCGAGGACGTG 1077

Qy   1167 GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC 1226
Db   1078 GCTCTCTACATCGGCCTTGTCGCTGTGGCTGTGTGCCTCTTCTTGCTGTTGCTGGCCCTT 1137

Qy   1227 ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1286
Db   1138 GGACTCATTTACTGTCGCAAGAAGGAAGGGCTGGACTCCGATGTGGCCGACTCGTCCATC 1197

Qy   1287 CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1346
Db   1198 CTCACCTCGGGCTTCCAGCCTGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCACCTG 1257

Qy   1347 CTCACCATCCAGCCGGACCTCAGCACCACC---ACCACCTACCAGGGCAGTCTCTGTCCC 1403
Db   1258 CTCACCATCCAGCCAGACCTCAGCACCACCACTACCACCTACCAGGGCAGTCTATGTTCG 1317

Qy   1404 CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG 1463
Db   1318 AGGCAGGATGGACCCAGCCCCAAGTTCCAGCTCTCTAATGGTCACCTGCTCAGCCCACTG 1377

Qy   1464 GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC 1523
Db   1378 GGGAGTGGCCGCCATACGTTGCACCACAGCTCACCCACCTCTGAGGCTGAGGACTTCGTC 1437

Qy   1524 TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC 1583
Db   1438 TCCCGCCTCTCCACCCAAAACTACTTTCGTTCCCTGCCCCGCGGCACCAGCAACATGGCC 1497

Qy   1584 TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC 1643
Db   1498 TACGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACGGGGATCAGCCTC 1557

Qy   1644 CTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC 1703
Db   1558 CTCATACCCCCGGATGCCATCCCCCGAGGAAAGATCTACGAGATCTACCTCACACTGCAC 1617

Qy   1704 AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT 1763
Db   1618 AAGCCAGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCAGTCGTT 1677

Qy   1764 AGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT 1823
Db   1678 AGCTGTGGGCCCCCAGGAGTCCTGCTCACCCGGCCAGTCATCCTTGCAATGGACCACTGT 1737

Qy   1824 GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC 1883
Db   1738 GGAGAGCCCAGCCCTGACAGCTGGAGTCTGCGCCTCAAAAAGCAGTCCTGCGAGGGCAGT 1797

Qy   1884 TGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG 1943
Db   1798 TGGG---AGGATGTGCTGCACCTTGGTGAGGAGTCACCTTCCCACCTCTACTACTGCCAG 1854

Qy   1944 CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA 2003
Db   1855 CTGGAGGCCGGGGCCTGCTATGTCTTCACGGAGCAGCTGGGCCGCTTTGCCCTGGTAGGA 1914

Qy   2004 GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC 2063
Db   1915 GAGGCCCTCAGCGTGGCTGCCACCAAGCGCCTCAGGCTCCTTCTGTTTGCTCCCGTGGCC 1974

Qy   2064 TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC 2123
Db   1975 TGTACGTCCCTTGAGTACAACATCCGAGTGTACTGCCTACACGACACCCACGACGCTCTC 2034

Qy   2124 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTC 2183
```

```
Db      2035 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTAGGTGGACAGCTGATCCAGGAGCCTCGCGTC 2094

Qy      2184 CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCC 2243
             ||||||||||| ||||||||||||||||| || || ||||||||||| ||||||||||||
Db      2095 CTGCACTTCAAAGACAGTTACCACAACCTACGTCTCTCCATCCACGACGTGCCCAGCTCC 2154

Qy      2244 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2303
             ||||||||||| ||||| ||||||||||||||||||||||| ||||| |||||||||||
Db      2155 CTGTGGAAGAGCAAGCTACTTGTCAGCTACCAGGAGATCCCTTTTTACCACATCTGGAAC 2214

Qy      2304 GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2363
             ||||| |||| |||  |||||||||||||||||||||||||  ||| | ||||||| ||
Db      2215 GGCACCCAGCAGTATCTGCACTGCACCTTCACCCTGGAGCGCATCAACGCCAGCACCAGC 2274

Qy      2364 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2423
             ||||||||||||||| |||||||||||||||||||||| || ||||||||||||| ||||
Db      2275 GACCTGGCCTGCAAGGTGTGGGTGTGGCAGGTGGAGGGAGATGGGCAGAGCTTCAACATC 2334

Qy      2424 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2483
             |||||||||||||| ||||||||||||||||||||  || |||||||||||||||||| |
Db      2335 AACTTCAACATCACTAAGGACACAAGGTTTGCTGAATTGTTGGCTCTGGAGAGTGAAGGG 2394

Qy      2484 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2543
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db      2395 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAAAAG 2454

Qy      2544 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2603
             || ||  |||| |||||||||||||| || ||||| ||||||||| | ||||| ||||||
Db      2455 ATCATCGCCAGTCTGGACCCACCCTGCAGCCGGGGCGCCGACTGGAGAACTCTAGCCCAG 2514

Qy      2604 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG 2663
             ||||| ||||||||||||||||| |||||||||||||||||||||||||| |||||||||
Db      2515 AAACTTCACCTGGACAGCCATCTTAGCTTCTTTGCCTCCAAGCCCAGCCCTACAGCCATG 2574

Qy      2664 ATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2723
             ||||||||||| |||||||| |||||||||||||||||||||||| |||||||||| |||
Db      2575 ATCCTCAACCTATGGGAGGCACGGCACTTCCCCAACGGCAACCTCGGCCAGCTGGCAGCA 2634

Qy      2724 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2783
             || ||||| ||||||||||| ||||| |||||||||||||| ||||||||||| |||||
Db      2635 GCTGTGGCCGGACTGGGCCAACCAGATGCTGGCCTCTTCACGGTGTCGGAGGCCGAGTGT 2694

Qy      2784 TGA 2786
             |||
Db      2695 TGA 2697


RESULT 9
BC058084
LOCUS       BC058084               3844 bp    mRNA    linear   ROD 21-OCT-2003
DEFINITION  Mus musculus unc-5 homolog A (C. elegans), mRNA (cDNA clone
            MGC:66671 IMAGE:6813463), complete cds.
ACCESSION   BC058084
VERSION     BC058084.1  GI:34784158
KEYWORDS    MGC.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 3844)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   MEDLINE  22388257
   PUBMED   12477932
REFERENCE   2  (bases 1 to 3844)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-SEP-2003) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Jim Lin, University of Iowa
            cDNA Library Preparation: M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Genome Sequence Centre,
            BC Cancer Agency, Vancouver, BC, Canada
            info@bcgsc.bc.ca
            Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
            Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
            Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
            Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
            Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
            Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
            Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
            George Yang, Scott Zuyderduyn, Marco Marra.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 126 Row: b Column: 11
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 23346570.
FEATURES             Location/Qualifiers
     source          1..3844
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="MGC:66671 IMAGE:6813463"
                     /tissue_type="Brain, mouse, 13.5, 14.5, 16.5, 17.5 dpc"
                     /clone_lib="NIH_BMAP_FW0"
                     /lab_host="DH10B"
                     /note="Vector: pYX-ASC"
     gene            1..3844
                     /gene="Unc5a"
                     /db_xref="LocusID:107448"
                     /db_xref="MGI:894682"
     CDS             252..2780
                     /codon_start=1
                     /product="Unc5a protein"
                     /protein_id="AAH58084.1"
                     /db_xref="GI:34784159"
                     /db_xref="LocusID:107448"
                     /translation="MAVRPGLWPALLGIVLTAWLRGSGAQQSATVANPVPGANPDLLP
                     HFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTM
                     EVRINVSRQQVEKVFGLEEYWCQCVAWSSSGTTKSQKAYIRIAYLRKNFEQEPLAKEV
                     SLEQGIVLPCRPPEGIPPAEVEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTPN
                     YTCVAKNIVARRRSASAAVIVYVDGSWSPWSKWSACGLDCTHWRSRECSDPAPRNGGE
                     ECRGADLDTRNCTSDLCLHTSSGPEDVALYIGLVAVAVCLILLLLVLVLIYCRKKEGL
                     DSDVADSSILTSGFQPVSIKPSKADNPHLLTIQPDLSTTTTTYQGSLCPRQDGPSPKF
                     QLSNGHLLSPLGSGRHTLHHSSPTSEAEDFVSRLSTQNYFRSLPRGTSNMAYGTFNFL
                     GGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPP
                     GVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGSWEDVLHLGEESPSHLYYCQLEAG
                     ACYVFTEQLGRFALVGEALSVAATKRLRLLLFAPVACTSLEYNIRVYCLHDTHDALKE
                     VVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWKSKLLVSYQEIPFYHIWN
```

```
                 GTQQYLHCTFTLERVNASTSDLACKVWVWQVEGDGQSFNINFNITKDTRFAEMLALES
                 EGGVPALVGPSAFKIPFLIRQKIITSLDPPCSRGADWRTLAQKLHLDSHLSFFASKPS
                 PTAMILNLWEARHFPNGNLGQLAAAVAGLGQPDAGLFTVSEAEC"
     misc_feature    714..926
                 /note="IG; Region: Immunoglobulin"
                 /db_xref="CDD:smart00409"
     misc_feature    984..1130
                 /note="TSP1; Region: Thrombospondin type 1 repeats"
                 /db_xref="CDD:smart00209"
     misc_feature    1566..1877
                 /note="ZU5; Region: Domain present in ZO-1 and Unc5-like
                 netrin receptors"
                 /db_xref="CDD:smart00218"
     misc_feature    2496..2684
                 /note="DEATH; Region: DEATH domain, found in proteins
                 involved in cell death (apoptosis). Alpha-helical domain
                 present in a variety of proteins with apoptotic functions.
                 Some (but not all) of these domains form homotypic and
                 heterotypic dimers"
                 /db_xref="CDD:smart00005"
ORIGIN

  Query Match            70.1%;  Score 2020.6;  DB 10;  Length 3844;
  Best Local Similarity  84.1%;  Pred. No. 0;
  Matches 2429;  Conservative    0;  Mismatches  274;  Indels  184;  Gaps    6;

Qy          1 AGCTGGGGCTCCGGGCTGAGGCGCT-AAAGCCGCCCTCCCGCCCGCGGGGCCCCGCGCCC 59
Db        169 AGGTGGGGCTCTCGGCTGAGGCGCTGACAGCCTCTCTCCCGCCCGCGGGGCCCCTAGTCC 228

Qy         60 GGCCCGCCCGCCTGCCCGCCCGCGGCCATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTC 119
Db        229 AGCCCGCTCGT----CCGCCCGCGGCCATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTC 284

Qy        120 CTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTG 179
Db        285 CTGGGCATAGTCCTCACTGCCTGGCTTCGTGGTTCGGGTGCCCAGCAGAGTGCCACAGTG 344

Qy        180 GCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAG 239
Db        345 GCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTGCCCCACTTCCTGGTAGAGCCGGAG 404

Qy        240 GATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACG 299
Db        405 GACGTGTACATTGTCAAGAACAAGCCCGTGCTGCTGGTGTGCAAGGCTGTGCCCGCCACC 464

Qy        300 CAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGC 359
Db        465 CAGATCTTCTTCAAGTGCAACGGGGAATGGGTTCGCCAGGTCGATCACGTCATTGAACGC 524

Qy        360 AGCACAGACGGGAGCAGTGGTGAGCCGACCATGGAGGTCCGCATTAATGTCTCAAGGCAG 419
Db        525 AGCACTGACGGCAGCAGCGGATTGCCAACCATGGAGGTCCGGATCAACGTATCAAGGCAG 584

Qy        420 CAGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCC 479
Db        585 CAGGTCGAGAAAGTGTTTGGGCTGGAGGAGTACTGGTGCCAGTGTGTGGCATGGAGCTCC 644

Qy        480 TCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCCAGATTGCGCAAGAACTTC 539
Db        645 TCAGGAACCACCAAAAGCCAGAAGGCCTACATCCGGATTGCCTATTTGCGCAAGAACTTT 704

Qy        540 GAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGT 599
Db        705 GAGCAGGAGCCGCTGGCCAAGGAAGTGTCACTGGAGCAAGGCATTGTGCTACCTTGTCGC 764

Qy        600 CCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGAC 659
Db        765 CCCCCGGAAGGAATCCCCCCAGCTGAGGTGGAGTGGCTCCGAAATGAGGACCTCGTGGAC 824

Qy        660 CCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCC 719
Db        825 CCCTCCCTCGACCCCAATGTGTACATCACACGGGAGCACAGCCTAGTCGTGCGGCAGGCC 884

Qy        720 CGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGC 779
Db        885 CGCCTGGCCGACACTCCCAACTACACCTGCGTGGCCAAGAACATCGTGGCCCGTCGCCGA 944

Qy        780 AGCGCCTCCGCTGCTGTCATCGTCTACGTGAACGGTGGGTGGTCGACGTGGACCGAGTGG 839
Db        945 AGCGCCTCTGCGGCCGTCA----------------------------------------- 963

Qy        840 TCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAAC 899
Db        964 ------------------------------------------------------------ 963

Qy        900 CCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAATGTCCATGACCGCACCGTC 959
Db        964 ------------------------------------------------------------ 963

Qy        960 TCCTCTCTGCTTGTCTCTGTGGACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGT 1019
Db        964 ---------TTGTTTATGTGGATGGGAGCTGGAGCCCATGGAGTAAGTGGTCAGCCTGC 1013

Qy       1020 GGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGA 1079
Db       1014 GGGCTTGACTGCACCCACTGGCGGAGCCGGGAGTGCTCCGACCCAGCGCCCCGCAACGGA 1073

Qy       1080 GGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGTACCAGTGACCTCTGTGTA 1139
Db       1074 GGTGAGGAGTGCCGGGGTGCTGACCTGGACACCCGCAACTGTACCAGTGACCTCTGCCTG 1133

Qy       1140 CACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTC 1199
Db       1134 CACACCTCTTCCGGCCCCGAGGACGTGGCTCTCTACATCGGCCTCGTCGCCGTGGCCGTG 1193

Qy       1200 TGCCTGGTCCTGCTGCTGCTTGTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTG 1259
Db       1194 TGCCTCATCTTGCTGCTGCTGGTCCTCGTCCTCATCTACTGCCGCAAGAAGGAAGGACTG 1253

Qy       1260 GACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAG 1319
Db       1254 GACTCAGACGTGGCTGACTCATCCATCCTTACCTCAGGCTTCCAGCCTGTCAGCATCAAG 1313

Qy       1320 CCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCGGACCTCAGCACCACC--- 1376
Db       1314 CCCAGCAAAGCAGACAATCCCCATCTGCTCACCATCCAACCGGACCTCAGCACCACCACG 1373

Qy       1377 ACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTC 1436
Db       1374 ACCACCTACCAGGGCAGCCTGTGTCCCCGGCAGGATGGACCCAGCCCCAAGTTCCAGCTC 1433

Qy       1437 ACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCACACACTGCACCACAGCTCT 1496
Db       1434 TCTAATGGTCACCTGCTCAGCCCACTGGGCAGTGGCCGCCATACGCTGCACCACAGCTCC 1493

Qy       1497 CCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCC 1556
Db       1494 CCCACCTCTGAGGCTGAGGACTTCGTCTCCCGCCTCTCCACCCAAAACTACTTTCGTTCT 1553

Qy       1557 CTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTG 1616
Db       1554 CTGCCCCGCGGTACCAGCAACATGGCCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTG 1613

Qy       1617 ATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCCAGATGCCATACCCCGAGGGAAG 1676
Db       1614 ATGATCCCTAACACAGGAATCAGCCTCCTCATACCCCCGGACGCCATCCCCCGAGGAAAG 1673

Qy       1677 ATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGC 1736
Db       1674 ATCTACGAGATCTACCTCACTCTGCACAAGCCAGAAGACGTGAGGTTGCCCCTAGCTGGC 1733

Qy       1737 TGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCTGGCGTCCTGCTCACCCGG 1796
Db       1734 TGTCAGACCCTGCTGAGTCCTATCGTTAGCTGTGGGCCCCCAGGAGTCCTGCTCACCCGG 1793
```

```
Qy   1797 CCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGC 1856
Db   1794 CCAGTCATCCTTGCCATGGACCACTGCGGGGAGCCCAGTCCCGACAGCTGGAGCCTGCGC 1853

Qy   1857 CTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGCAGGATGTGCTGCACCTGGGCGAGGAG 1916
Db   1854 CTCAAAAAGCAGTCCTGTGAGGGCAGCTGGG---AGGACGTGCTGCACCTTGGTGAGGAG 1910

Qy   1917 GCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAG 1976
Db   1911 TCGCCCTCTCATCTCTACTACTGCCAGCTGGAGGCCGGGGCCTGCTATGTCTTTACCGAG 1970

Qy   1977 CAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTC 2036
Db   1971 CAGCTAGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCACCAAGCGCCTC 2030

Qy   2037 AAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTAC 2096
Db   2031 AGGCTCCTTCTGTTTGCCCCTGTGGCCTGTACGTCCCTCGAGTACAACATCCGAGTGTAC 2090

Qy   2097 TGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGG 2156
Db   2091 TGCCTGCACGACACCCACGATGCTCTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGT 2150

Qy   2157 GGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGC 2216
Db   2151 GGACAGCTGATCCAGGAGCCCCGTGTCCTGCACTTCAAAGACAGTTACCACAACCTACGT 2210

Qy   2217 CTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAG 2276
Db   2211 CTGTCCATCCACGACGTGCCCAGCTCCCTGTGGAAGAGCAAGCTCCTTGTCAGCTACCAG 2270

Qy   2277 GAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTTGCACTGCACCTTCACC 2336
Db   2271 GAGATCCCTTTTTACCACATCTGGAATGGCACTCAGCAGTATCTGCACTGCACCTTCACC 2330

Qy   2337 CTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTG 2396
Db   2331 CTGGAGCGCGTCAATGCCAGCACCAGCGACCTGGCCTGCAAGGTGTGGGTGTGGCAGGTG 2390

Qy   2397 GAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACCAAGGACACAAGGTTTGCT 2456
Db   2391 GAGGGAGATGGACAGAGCTTCAACATCAACTTTAACATCACTAAGGACACGAGGTTTGCT 2450

Qy   2457 GAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTC 2516
Db   2451 GAAATGCTGGCTCTGGAGAGTGAAGGGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTC 2510

Qy   2517 AAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGG 2576
Db   2511 AAGATCCCCTTCCTCATTCGGCAAAAGATCATTACCAGCCTGGACCCACCCTGCAGCCGG 2570

Qy   2577 GGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTT 2636
Db   2571 GGCGCCGACTGGCGAACTCTAGCCCAGAAACTTCACCTGGACAGCCATCTTAGCTTCTTT 2630

Qy   2637 GCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCC 2696
Db   2631 GCCTCCAAGCCCAGCCCTACAGCCATGATCCTCAACCTATGGGAGGCGCGGCACTTCCCC 2690

Qy   2697 AACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGC 2756
Db   2691 AACGGCAACCTCGGCCAGCTGGCCGCAGCTGTGGCCGGACTGGGCCAGCCAGATGCTGGC 2750

Qy   2757 CTCTTCACAGTGTCGGAGGCTGAGTGCTGAGGCCGGCCAGGCCCGACA--CCTACACTCT 2814
Db   2751 CTCTTCACCGTGTCAGAGGCCGAGTGCTGAGACCAGCCAGGCCGGTCACGCCTACACTCT 2810

Qy   2815 CACCAGCTTTGGCACCCACCAAGGACAGGCAGAAGCCGGACAGGGGCCCTTCCCCACACC 2874
Db   2811 CACCAGCTTTGGCACCTGCCAGGGACAGGCAAAAGCCAGACAGGGGCCCTACCCCCACAC 2870

Qy   2875 GGGGAGA 2881

Db   2871 CCGGGGA 2877
```

```
RESULT 10
BC009333
LOCUS       BC009333                 2688 bp    mRNA    linear   PRI 19-DEC-2003
DEFINITION  Homo sapiens cDNA clone IMAGE:4126760, partial cds.
ACCESSION   BC009333
VERSION     BC009333.2  GI:40226527
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2688)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
REFERENCE   2  (bases 1 to 2688)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JUN-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     On Dec 19, 2003 this sequence version replaced gi:14424611.
            Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC),
            Gaithersburg, Maryland;
            Web site: http://www.nisc.nih.gov/
            Contact: nisc_mgc@nhgri.nih.gov
            Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
            Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
            Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
            Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
            Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D.,McCloskey,J.C.,
            McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
            Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
            Young,A., Zhang,L.-H. and Green,E.D.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAL Plate: 26 Row: g Column: 22.
FEATURES             Location/Qualifiers
     source          1..2688
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
```

```
                     /clone="IMAGE:4126760"
                     /tissue_type="Brain, neuroblastoma"
                     /clone_lib="NIH_MGC_19"
                     /lab_host="DH10B-R"
                     /note="Vector: pOTB7"
     CDS             <1..1627
                     /codon_start=2
                     /product="Unknown (protein for IMAGE:4126760)"
                     /protein_id="AAH09333.2"
                     /db_xref="GI:40226528"
                     /translation="DVALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILT
                     SGFQPVSIKPSKADNPHLLTIQPDLSTTTTTYQGSLCPRQDGPSPKFQLTNGHLLSPL
                     GGGRHTLHHSSPTSEAEEFVSRLSTQNYFRSLPRGTSNMTYGTFNFLGGRLMIPNTGI
                     SLLIPPDAIPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILA
                     MDHCGEPSPDSWSLRLKKQSCEGSWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGR
                     FALVGEALSVAAAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQ
                     LIQEPRVLHFKDSYHNLRLSIHDVPSSLWKSKLLVSYQEIPFYHIWNGTQRYLHCTFT
                     LERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPS
                     AFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLDSHLSFFASKPSPTAMILNLWEA
                     RHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC"
     misc_feature    413..724
                     /note="ZU5; Region: Domain present in ZO-1 and Unc5-like
                     netrin receptors"
                     /db_xref="CDD:smart00218"
     misc_feature    1343..1594
                     /note="DEATH; Region: DEATH domain, found in proteins
                     involved in cell death (apoptosis)"
                     /db_xref="CDD:smart00005"
ORIGIN

  Query Match            58.7%;  Score 1691.4;  DB 9;  Length 2688;
  Best Local Similarity  99.6%;  Pred. No. 7e-270;
  Matches 1718;  Conservative    0;  Mismatches    1;  Indels    6;  Gaps    2;

Qy       1160 GGACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCT 1219
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCT 60

Qy       1220 TGTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTC 1279
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTC 120

Qy       1280 GTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCC 1339
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCC 180

Qy       1340 CCATCTGCTCACCATCCAGCCGGACCTCAG---CACCACCACCACCTACCAGGGCAGTCT 1396
              ||||||||||||||||||||||||||||||   |||||||||||||||||||||||||||
Db        181 CCATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCT 240

Qy       1397 CTGTCCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAG 1456
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTGTCCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAG 300

Qy       1457 CCCCCTGGGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGA 1516
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCCCCTGGGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGA 360

Qy       1517 GTTCGTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAA 1576
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GTTCGTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAA 420

Qy       1577 CATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTAT 1636
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        421 CATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGAAT 480

Qy       1637 CAGCCTCCTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCAC 1696
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CAGCCTCCTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCAC 540

Qy       1697 GCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCC 1756
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCC 600

Qy       1757 CATCGTTAGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGA 1816
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CATCGTTAGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGA 660

Qy       1817 CCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGA 1876
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGA 720

Qy       1877 GGGCAGCTGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTA 1936
              |||||||||||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GGGCAGCTGGG---AGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTA 777

Qy       1937 CTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCT 1996
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        778 CTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCT 837

Qy       1997 GGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCC 2056
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        838 GGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCC 897

Qy       2057 GGTGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGA 2116
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        898 GGTGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGA 957

Qy       2117 TGCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCC 2176
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        958 TGCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCC 1017

Qy       2177 ACGGGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCC 2236
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1018 ACGGGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCC 1077

Qy       2237 CAGCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACAT 2296
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1078 CAGCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACAT 1137

Qy       2297 CTGGAATGGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAG 2356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1138 CTGGAATGGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAG 1197

Qy       2357 CACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTT 2416
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1198 CACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTT 1257

Qy       2417 CAGCATCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAG 2476
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1258 CAGCATCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAG 1317

Qy       2477 TGAAGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCG 2536
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1318 TGAAGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCG 1377

Qy       2537 GCAGAAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCT 2596
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1378 GCAGAAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCT 1437

Qy       2597 GGCCCAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCAC 2656
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1438 GGCCCAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCAC 1497

Qy       2657 AGCCATGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCT 2716
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1498 AGCCATGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCT 1557

Qy       2717 GGCTGCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGC 2776
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1558 GGCTGCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGC 1617

Qy       2777 TGAGTGCTGAGGCCGGCCAGGCCCGACACCTACACTCTCACCAGCTTTGGCACCCACCAA 2836
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1618 TGAGTGCTGAGGCCGGCCAGGCCCGACACCTACACTCTCACCAGCTTTGGCACCCACCAA 1677
```

```
Qy      2837 GGACAGGCAGAAGCCGGACAGGGGCCCTTCCCCACACCGGGGAGA 2881
Db      1678 GGACAGGCAGAAGCCGGACAGGGGCCCTTCCCCACACCGGGGAGA 1722

RESULT 11
BD057525
LOCUS       BD057525                1787 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Netrin receptors.
ACCESSION   BD057525
VERSION     BD057525.1  GI:22603131
KEYWORDS    JP 2001505062-A/2.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 1787)
  AUTHORS   Lavigne,M.T., Leonardo,D.E., Hinck,L., Masu,M. and Masu,K.K.
  TITLE     Netrin receptors
  JOURNAL   Patent: JP 2001505062-A 2 17-APR-2001;
            THE REGENTS OF THE UNIV OF CALIFORNIA
COMMENT     PN   JP 2001505062-A/2
            PD   17-APR-2001
            PF   19-FEB-1998 JP 1998536840
            PR   19-FEB-1997 US     08/808982
            PI   MARC TESSIER LAVIGNE,DAVID E LEONARDO,LINDSAY HINCK,MASAYUKI
            PI   MASU,
            PI   KAZUKO KEINO MASU
            PC   C07K1/00,C07K14/00,C07K17/00,C07H21/02,C07H21/04,G01N33/53 CC
             Strandedness: Double;
            CC   Topology: Linear;
            FH   Key             Location/Qualifiers.
FEATURES             Location/Qualifiers
     source          1..1787
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match            54.5%;  Score 1570.4;  DB 6;  Length 1787;
  Best Local Similarity  96.9%;  Pred. No. 7.5e-250;
  Matches 1721;  Conservative    0;  Mismatches   36;  Indels   19;  Gaps   11;

Qy      1114 GCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCT 1173
Db         1 GCAACTGTACCAGTGACCTCTG-GTACACACTGCTTCTGGCCCTGAGGACGTGGCCCTCT 59

Qy      1174 ATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCATCCTCG 1233
Db        60 ATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCATCCTCG 119

Qy      1234 TTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCT 1293
Db       120 TTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCT 179

Qy      1294 CAGGCTTCCAGCCCGTCAGCATC-AAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACC 1352
Db       180 CAGGCTTCCAGCCCGTCAGCATCTAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACC 239

Qy      1353 ATCCAGCCGGACCTCAG---CACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAG 1409
Db       240 ATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAG 299

Qy      1410 GATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGC 1469
Db       300 GATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGC 359

Qy      1470 GGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGC 1529
Db       360 GGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGC 419

Qy      1530 CTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGG 1589
Db       420 CTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGG 479

Qy      1590 ACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATC 1649
Db       480 ACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGAATCAGCCTCCTCATC 539

Qy      1650 CCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCG 1709
Db       540 CCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCG 599

Qy      1710 GAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGT 1769
Db       600 GAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGT 659

Qy      1770 GGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAG 1829
Db       660 GGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAG 719

Qy      1830 CCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAG 1889
Db       720 CCCAGCCCTGACAGCTGGAGCCTGGCCCTCAAAAAGCAGTCGTGCGAGGG-AGCTGGGAG 778

Qy      1890 CAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG 1949
Db       779 GATG----TCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG 834

Qy      1950 GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC 2009
Db       835 GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC 894

Qy      2010 CTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC 2069
Db       895 CTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC 954

Qy      2070 TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG 2129
Db       955 TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG 1014

Qy      2130 GTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC 2189
Db      1015 GTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC 1074

Qy      2190 TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG 2249
Db      1075 TT-AAGGACAGTTACCACAACCT--GCCCTATCATCCACGATGTGCCCAGCTCCCTGTGG 1131

Qy      2250 AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG 2309
Db      1132 AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG 1191

Qy      2310 CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG 2369
Db      1192 CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG 1251

Qy      2370 GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC 2429
Db      1252 GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC 1311

Qy      2430 AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC 2489
Db      1312 AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC 1371

Qy      2490 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT 2549
Db      1372 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT 1431

Qy      2550 TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC 2609
Db      1432 TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC 1491

Qy      2610 CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC 2669
Db      1492 CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC 1551
```

```
Qy       2670 AACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG 2729
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1552 AACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG 1611

Qy       2730 GCTGGACTGGGCCAGCCAGACGCTGGCCTC-TTCACAGTG-TCGGAGGCTGAGTGCTGAG 2787
              |||||     |||||||  |||| |||| || |||||||| ||||||||||||||||||
Db       1612 GCTGGGACTGGCCAGCAGGACGGTGGCTTCTTTCACAGTGTTCGGAGGCTGAGTGCTGAG 1671

Qy       2788 GCCGGCCAGGCCCGACACCTACACTCTCACCAGCTTTGGCA--CCCACCAAGGACAGGCA 2845
              |||||||||  |||   ||||| | | ||||| |||||  |   ||  | | ||||||||
Db       1672 GCCGGCCAGG--CGAACACTACAATTTTACCAGTTTTGGGAACCCACCAAGGGACAGGCA 1729

Qy       2846 GAAGCCGGACAGGGGCCCTTCCCCACACCGGGGAGA 2881
              |||||||||| ||||  ||||| | ||||||||||
Db       1730 GAAGCCGGACAAGGGCTTTTCCCAAAACCGGGGAGA 1765
```

RESULT 12
AX054976
LOCUS       AX054976                9700 bp    DNA     linear   PAT 13-JAN-2001
DEFINITION  Sequence 91 from Patent WO0073328.
ACCESSION   AX054976
VERSION     AX054976.1  GI:12228344
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   van Criekinge,W., Roelens,I., Bogaert,T. and Verwaerde,P.
  TITLE     Unc-5 constructs and screening methods
  JOURNAL   Patent: WO 0073328-A 91 07-DEC-2000;
            Devgen NV (BE)
FEATURES             Location/Qualifiers
     source          1..9700
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="plasmid pGC1037"
ORIGIN

  Query Match             44.8%;  Score 1291.2;  DB 6;  Length 9700;
  Best Local Similarity   99.5%;  Pred. No. 9e-204;
  Matches 1306;  Conservative    0;  Mismatches    3;  Indels    3;  Gaps    1;

```
Qy       1478 CACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCAC 1537
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 CACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCAC 66

Qy       1538 CCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAA 1597
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 CCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAA 126

Qy       1598 CTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCCAGA 1657
              |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db        127 CTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGAATCAGCCTCCTCATCCCCCCAGA 186

Qy       1658 TGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGT 1717
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 TGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGT 246

Qy       1718 GAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCC 1777
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 GAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCC 306

Qy       1778 TGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCC 1837
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 TGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCC 366

Qy       1838 TGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGCAGGATGT 1897
              |||||||||||||||||||||||||||||||||||||||||||||||||   |||||||
Db        367 TGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGG---AGGATGT 423

Qy       1898 GCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGC 1957
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 GCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGC 483

Qy       1958 CTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGT 2017
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 CTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGT 543

Qy       2018 GGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGA 2077
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        544 GGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGA 603

Qy       2078 GTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCA 2137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        604 GTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCA 663

Qy       2138 GCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGA 2197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        664 GCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGA 723

Qy       2198 CAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAA 2257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        724 CAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAA 783

Qy       2258 GCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTA 2317
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        784 GCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTA 843

Qy       2318 CTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAA 2377
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        844 CTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAA 903

Qy       2378 GCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCAC 2437
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        904 GCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCAC 963

Qy       2438 CAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCT 2497
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        964 CAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCAAGCCCT 1023

Qy       2498 GGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCT 2557
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1024 GGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCT 1083

Qy       2558 GGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGA 2617
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1084 GGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGA 1143

Qy       2618 CAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTG 2677
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1144 CAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTG 1203

Qy       2678 GGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACT 2737
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1204 GGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACT 1263

Qy       2738 GGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGAGGC 2789
              |||||||||||||||||||||||||||||||||||||||||||||||  |
Db       1264 GGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGAGTC 1315
```

RESULT 13
MMU72634
LOCUS       MMU72634                9299 bp    mRNA    linear   ROD 13-MAY-1997
DEFINITION  Mus musculus rostral cerebellar malformation protein (rcm) mRNA,
            complete cds.
ACCESSION   U72634
VERSION     U72634.1  GI:2088526
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 9299)

```
AUTHORS   Ackerman,S.L., Kozak,L.P., Przyborski,S.A., Rund,L.A., Boyer,B.B.
          and Knowles,B.B.
TITLE     The mouse rostral cerebellar malformation gene encodes an
          UNC-5-like protein
JOURNAL   Nature 386 (6627), 838-842 (1997)
MEDLINE   97271898
PUBMED    9126743
REFERENCE 2  (bases 1 to 9299)
AUTHORS   Ackerman,S.L., Kozak,L.P., Rund,L.A. and Knowles,B.B.
TITLE     Direct Submission
JOURNAL   Submitted (25-SEP-1996) The Jackson Laboratory, 600 Main Street,
          Bar Harbor, ME 04609, USA
FEATURES             Location/Qualifiers
     source          1..9299
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57B6/SJL"
                     /db_xref="taxon:10090"
                     /chromosome="3"
                     /map="between Amy1 and Mpmv9"
     gene            1..9299
                     /gene="rcm"
     CDS             151..2946
                     /gene="rcm"
                     /function="transmembrane receptor"
                     /note="similar to C. elegans UNC-5"
                     /codon_start=1
                     /product="rostral cerebellar malformation protein"
                     /protein_id="AAB54103.1"
                     /db_xref="GI:2088527"
                     /translation="MRKGLRATAARCGLGLGYLLQMLVLPALALLSASGTGSAAQDDE
                     FFHELPETFPSDPPEPLPHFLIEPEEAYIVKNKPVNLYCKASPATQIYFKCNSEWVHQ
                     KDHVVDERVDETSGLIVREVSIEISRQQVEELFGPEDYWCQCVAWSSAGTTKSRKAYV
                     RIAYLRKTFEQEPLGKEVSLEQEVLLQCRPPEGIPVAEVEWLKNEDIIDPAEDRNFYI
                     TIDHNLIIKQARLSDTANYTCVAKNIVAKRKSTTATVIVYVNGGWSTWTEWSVCNSRC
                     GRGYQKRTRTCTNPAPLNGGAFCEGQSVQKIACTTLCPVDGRWTSWSKWSTCGTECTH
                     WRRRECTAPAPKNGGKDCDGLVLQSKNCTDGLCMQAAPDSDDVALYVGIVIAVTVCLA
                     ITVVVALFVYRKNHRDFESDIIDSSALNGGFQPVNIKAARQDLLAVPPDLTSAAAMYR
                     GPVYALHDVSDKIPMTNSPILDPLPNLKIKVYNSSGAVTPQDDLAEFSSKLSPQMTQS
                     LLENEALNLKNQSLARQTDPSCTAFGTFNSLGGHLIIPNSGVSLLIPAGAIPQGRVYE
                     MYVTVHRKENMRPPMEDSQTLLTPVVSCGPPGALLTRPVILTLHHCADPSTEDWKIQL
                     KNQAVQGQWEDVVVVGEENFTTPCYIQLDAEACHILTENLSTYALVGQSTTKAAAKRL
                     KLAIFGPLCCSSLEYSIRVYCLDDTQDALKEVLQLERQMGGQLLEEPKALHFKGSIHN
                     LRLSIHDIAHSLWKSKLLAKYQEIPFYHIWSGSQRNLHCTFTLERLSLNTVELVCKLC
                     VRQVEGEGQIFQLNCTVSEEPTGIDLPLLDPASTITTVTGPSAFSIPLPIRQKLCSSL
                     DAPQTRGHDWRMLAHKLNLDRYLNYFATKSSPTGVILDLWEAQNFPDGNLSMLAAVLE
                     EMGRHETVVSLAAEGQY"
ORIGIN

  Query Match            32.6%;  Score 940.2;  DB 10;  Length 9299;
  Best Local Similarity  61.9%;  Pred. No. 9.9e-146;
  Matches 1642;  Conservative    0;  Mismatches  943;  Indels   66;  Gaps    7;

Qy        190 TGCCTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACA 249
Db        308 TTCCATCTGACCCACCTGAGCCATTGCCACACTTCCTCATTGAGCCCGAGGAAGCTTACA 367

Qy        250 TCGTCAAGAACAAGCCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCT 309
Db        368 TTGTGAAGAACAAGCCTGTGAACCTGTATTGTAAAGCCAGCCCTGCCACCCAGATCTACT 427

Qy        310 TCAAGTGCAACGGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACG 369
Db        428 TCAAGTGCAACAGCGAGTGGGTTCATCAGAAGGACCACGTAGTAGACGAGAGAGTAGATG 487

Qy        370 GGAGCAGTGGTGAGCCGACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGA 429
Db        488 AAACCTCTGGTCTAATTGTGAGAGAAGTGAGCATTGAGATTTCACGCCAGCAGGTGGAGG 547

Qy        430 AGGTGTTCGGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCA 489
Db        548 AACTGTTTGGGCCTGAAGATTACTGGTGCCAGTGTGTGGCCTGGAGCTCAGCAGGCACTA 607

Qy        490 CCAAGAGTCAGAAGGCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGC 549
Db        608 CGAAGAGTCGGAAGGCATACGTGCGCATTGCGTATCTGCGGAAGACATTCGAGCAGGAAC 667

Qy        550 CGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGG 609
Db        668 CCTTGGGAAAGGAAGTGTCCTTGGAGCAGGAAGTCTTACTCCAGTGTCGGCCACCTGAAG 727

Qy        610 GCATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGG 669
Db        728 GGATCCCAGTGGCTGAGGTGGAATGGCTAAAGAATGAAGACATAATTGATCCTGCTGAAG 787

Qy        670 ACCCCAATGTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTG 729
Db        788 ATCGGAACTTTTATATTACTATCGATCACAACCTGATCATCAAGCAAGCCCGACTCTCAG 847

Qy        730 ACACGGCCAACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCG 789
Db        848 ATACAGCAAATTATACCTGTGTTGCCAAAAATATTGTTGCCAAGAGAAAAAGCACCACAG 907

Qy        790 CTGCTGTCATCGTCTACGTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCA 849
Db        908 CCACTGTCATCGTGTATGTTAATGGTGGCTGGTCCACCTGGACAGAGTGGTCTGTGTGTA 967

Qy        850 GCGCCAGCTGTGGGCGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTC 909
Db        968 ACAGCCGCTGTGGGCGAGGATATCAGAAACGCACAAGAACCTGCACCAACCCAGCCCCAC 1027

Qy        910 TCAACGGGGGCGCTTTCTGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCCTCTCTGC 969
Db       1028 TCAATGGTGGGGCCTTCTGTGAGGGGCAGAGTGTGCAGAA---AATAGCATGCACTACGT 1084

Qy        970 TTGTCTCTGTGGACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACT 1029
Db       1085 TATGTCCAGTGGATGGTAGGTGGACTTCATGGAGCAAATGGTCAACCTGTGGGACTGAAT 1144

Qy       1030 GCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGT 1089
Db       1145 GCACCCACTGGCGCAGGAGGGAGTGTACAGCACCAGCCCCCAAGAACGGGGGTAAGGACT 1204

Qy       1090 GCCAGGGCACTGACCTGGACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTT 1149
Db       1205 GTGATGGCCTGGTCCTCCAATCCAAGAACTGCACTGATGGGCTGTGCATGCAGGCTGCTC 1264

Qy       1150 CTGGCCCTGAGGACGTGGCCCTCTATGTGGG---CCTCATCGCCGTGGCCGTCTGCCTGG 1206
Db       1265 CTGACTCAGATGATGTGGCTCTCTACGTGGGGATTGTGATCGCTGTAACAGTCTGTCTGG 1324

Qy       1207 TCCTGCTGCTGCTTGTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAG 1266
Db       1325 CGATCACTGTTGTGGTGGCCCTGTTTGTGTATCGGAAGAACCACCGTGACTTTGAGTCTG 1384

Qy       1267 ATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCA 1326
Db       1385 ACATCATTGACTCCTCAGCACTCAATGGCGGCTTTCAGCCTGTGAACATCAAGGCTGCCA 1444

Qy       1327 AAGCAGACAACCCCCATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCACCTACC 1386
Db       1445 GACAAGATCTCCTGGCTGTCCCCCCTGACC------TCACCTCAGCTGCAGCCATGTACA 1498

Qy       1387 AGGGCAGTCTCTGTCCCCGGCAGGATG-----GGCCCAGCCCCAAGTTCCAGCTCACCAA 1441
Db       1499 GGGGACCTGTCTATGCTCTGCATGATGTCTCAGACAAAATCCCAATGACCAACTCTCCAA 1558

Qy       1442 TGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCAC 1501
Db       1559 TTCTGGACCCACTACCCAACTTGAAAATCAAAGTGTACAACAGCTCAGGTGCTGTCACTC 1618

Qy       1502 CTC----TGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACCC----------------- 1539
Db       1619 CTCAGGATGACCTTGCCGAGTTCTCATCCAAACTGTCACCCCAGATGACCCAGTCCTTGC 1678

Qy       1540 ------------------------AGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCA 1575
```

```
Db   1679 TAGAGAATGAGGCCCTTAACCTGAAGAACCAGAGCCTCGCAAGACAGACTGACCCATCCT 1738

Qy   1576 ACATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTA 1635

Db   1739 GCACAGCATTTGGTACCTTCAACTCTCTTGGGGGTCACCTCATCATTCCTAATTCAGGAG 1798

Qy   1636 TCAGCCTCCTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCA 1695

Db   1799 TAAGCTTGCTGATTCCCGCTGGGGCCATTCCTCAGGGGAGAGTCTATGAAATGTATGTGA 1858

Qy   1696 CGCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTC 1755

Db   1859 CTGTACACAGGAAAGAAAATATGAGGCCCCCCATGGAAGACTCTCAGACCCTACTTACCC 1918

Qy   1756 CCATCGTTAGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGG 1815

Db   1919 CTGTGGTGAGCTGTGGGCCTCCTGGAGCTCTGCTGACCCGCCCTGTCATCCTCACTCTGC 1978

Qy   1816 ACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCG 1875

Db   1979 ATCACTGTGCAGACCCCAGCACCGAGGACTGGAAGATCCAGCTCAAAAACCAGGCAGTGC 2038

Qy   1876 AGGGCAGCTGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACT 1935

Db   2039 AGGGACAATGGGAGGATGTTGTGGTG---GTTGGGGAGGAGAACTTCACAACCCCCTGTT 2095

Qy   1936 ACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCC 1995

Db   2096 ACATTCAGCTGGATGCAGAGGCTTGCCATATCCTCACAGAGAACCTCAGTACCTATGCCC 2155

Qy   1996 TGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGC 2055

Db   2156 TGGTTGGGCAGTCCACCACCAAAGCAGCTGCCAAGCGTCTTAAACTGGCCATCTTTGGGC 2215

Qy   2056 CGGTGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACG 2115

Db   2216 CCCTCTGCTGCTCTTCCCTGGAGTACAGCATTAGAGTCTACTGCCTGGATGACACACAGG 2275

Qy   2116 ATGCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGC 2175

Db   2276 ATGCCCTGAAGGAAGTTCTACAACTGGAGAGGCAAATGGGAGGACAGCTCCTAGAAGAAC 2335

Qy   2176 CACGGGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGC 2235

Db   2336 CCAAGGCTCTTCATTTTAAAGGCAGCATCCACAACCTGCGCCTGTCTATTCATGACATCG 2395

Qy   2236 CCAGCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACA 2295

Db   2396 CCCATTCCCTCTGGAAGAGCAAATTGCTGGCTAAGTATCAGGAAATTCCATTTTACCACA 2455

Qy   2296 TCTGGAATGGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCA 2355

Db   2456 TCTGGAGTGGCTCTCAAAGAAACCTCCACTGCACCTTCACTCTGGAAAGACTCAGCCTAA 2515

Qy   2356 GCACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCT 2415

Db   2516 ACACAGTGGAACTGGTTTGCAAACTCTGTGTGCGGCAGGTTGAAGGAGAAGGGCAGATCT 2575

Qy   2416 TCAGCATCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGA 2475

Db   2576 TCCAGCTCAACTGTACTGTGTCAGAGGAACCTACTGGCATCGACTTACCTCTCCTGGACC 2635

Qy   2476 GTGAAGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTC 2535

Db   2636 CTGCTAGTACCATCACCACTGTCACCGGACCAAGTGCTTTCAGCATTCCTCTCCCTATCC 2695

Qy   2536 GGCAGAAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTC 2595

Db   2696 GGCAGAAGCTATGCAGCAGCCTGGATGCCCCTCAAACAAGAGGCCATGACTGGAGGATGC 2755

Qy   2596 TGGCCCAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCA 2655

Db   2756 TGGCCCATAAACTCAACCTGGACAGGTACTTGAATTACTTTGCCACCAAATCGAGCCCAA 2815

Qy   2656 CAGCCATGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGC 2715

Db   2816 CTGGCGTAATCCTGGATCTTTGGGAAGCACAGAACTTCCCAGATGGAAACCTGAGCATGC 2875

Qy   2716 TGGCTGCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGG 2775

Db   2876 TGGCAGCCGTCCTGGAAGAAATGGGAAGACATGAGACAGTGGTGTCCTTGGCAGCAGAAG 2935

Qy   2776 CTGAGTGCTGA 2786

Db   2936 GACAGTATTGA 2946
```

```
RESULT 14
AY187310
LOCUS       AY187310                2962 bp    mRNA    linear   VRT 06-JUN-2003
DEFINITION  Gallus gallus UNC5-like protein 3 mRNA, complete cds.
ACCESSION   AY187310
VERSION     AY187310.1  GI:31442350
KEYWORDS    .
SOURCE      Gallus gallus (chicken)
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 2962)
  AUTHORS   Guan,W. and Condic,M.L.
  TITLE     Characterization of Netrin-1, Neogenin and cUNC-5H3 expression
            during chick dorsal root ganglia development
  JOURNAL   Gene Expr. Patterns 3, 369-373 (2003)
REFERENCE   2  (bases 1 to 2962)
  AUTHORS   Guan,W. and Condic,M.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-NOV-2002) Neurobiology & Anatomy, University of Utah,
            20 North, 1900 East, Salt Lake City, UT 84132-3401, USA
FEATURES             Location/Qualifiers
     source          1..2962
                     /organism="Gallus gallus"
                     /mol_type="mRNA"
                     /db_xref="taxon:9031"
     CDS             1..2796
                     /codon_start=1
                     /product="UNC5-like protein 3"
                     /protein_id="AAO67275.1"
                     /db_xref="GI:31442351"
                     /translation="MGKGLEGTAARCGLGMGYLLHSVVLPALAVLGASRPGSAAQDDD
                     FFHELPETFPSDPPEPLPHFLIEPEEAYIVKNKPVNLYCKASPATQIYFKCNSEWVHQ
                     KDHVVDERVDETSGLIVCEVSIEISRQQVEELFGPEDYWCQCVAWSSAGTTKSRKAYV
                     RIAYLRKTFEQEPLGKEVSLEQEVLLQCRPPEGIPVAEVEWLKNEEVIDPVEDRNFYI
                     TIDHNLIIKQARLSDTANYTCVAKNIVAKRKSTTATVIVYVNGGWSTWTEWSACNSRC
                     GRGFQKRTRTCTNPAPLNGGAFCEGQNVQKIACTTLCPVDGKWTSWSKWSTCGTECTH
                     WRRRECTAPAPKNGGKDCEGLVLQSKNCTDGLCMQAAPDSDDVALYVGIVIAVIVCLA
                     ISVVVALFVYRKNHRDFESDIIDSSALNGGFQPVNIKAARQDLLAVPPDLTSAAAMYR
                     GPVYALHDVSDKIPMTNSPILDPLPNLKIKVYNTSGAVTPQDELSDFSSKLSPQITQS
                     LLENETLNVKNQSLARQTDPSCTAFGTFNSLGGHLVIPNSGVSLLIPAGAVPQGRVYE
                     MYVTVHRKEGMRPPVEDSQTLLTPVVSCGPPGALLTRPVVLTMHHCAEPNMDDWQIQL
                     KHQAGQGPWEDVVVVGEENFTTPCYIQLDPEACHILTETLSTYALVGQSITKAAAKRL
                     KLAIFGPLSCSSLEYSIRVYCLDDTQDALKEVLQLERQMGGQLLEEPKTLHFKGSTHN
                     LRLSIHDIAHSLWKSKLPAKYQEIPFYHIWSGCQRNLHCTFTLERFSLNTLELVCKLC
                     VRQVEGEGQIFQLNCSVSEEPTGIDYPIMDSAGSITTIVGPNAFSIPLPIRQKLCSSL
                     DAPQTRGHDWRMLAHKLKLDRYLNYFATKSSPTGVILDLWEAQNFPDGNLSMLAAVLE
                     EMGRHETVVSLAAEGNY"
ORIGIN

  Query Match             32.6%;  Score 938.4;  DB 5;  Length 2962;
  Best Local Similarity   61.3%;  Pred. No. 2.2e-145;
  Matches 1637;  Conservative    0;  Mismatches  961;  Indels   72;  Gaps    5;

Qy        190 TGCCTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACA 249

Db        158 TTCCTTCTGATCCTCCAGAGCCATTGCCCCACTTTCTCATTGAACCCGAAGAAGCTTACA 217
```

```
Qy    250 TCGTCAAGAACAAGCCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCT 309
Db    218 TCGTGAAAAACAAGCCTGTGAATCTGTACTGCAAAGCGAGCCCTGCCACGCAGATCTATT 277

Qy    310 TCAAGTGCAACGGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACG 369
Db    278 TTAAGTGCAACAGTGAATGGGTTCATCAGAAGGATCATGTGGTGGATGAGAGAGTAGATG 337

Qy    370 GGAGCAGTGGTGAGCCGACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGA 429
Db    338 AAACCTCTGGTCTGATCGTCTGTGAGGTGAGCATCGAGATTTCCCGCCAGCAGGTGGAAG 397

Qy    430 AGGTGTTCGGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCA 489
Db    398 AGCTCTTTGGACCCGAGGACTACTGGTGCCAGTGTGTCGCCTGGAGCTCAGCTGGCACCA 457

Qy    490 CCAAGAGTCAGAAGGCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGC 549
Db    458 CCAAGAGCCGCAAGGCCTACGTCCGCATTGCATATCTCAGAAAGACTTTTGAGCAGGAGC 517

Qy    550 CGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGG 609
Db    518 CGCTGGGGAAAGAAGTGTCCCTGGAGCAAGAGGTCCTGCTCCAGTGCCGTCCTCCTGAAG 577

Qy    610 GCATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGG 669
Db    578 GCATTCCAGTAGCTGAGGTAGAGTGGCTGAAGAATGAAGAGGTGATCGATCCTGTGGAAG 637

Qy    670 ACCCCAATGTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTG 729
Db    638 ACCGAAATTTTTACATCACCATTGATCACAACCTGATCATCAAGCAAGCCCGGCTTTCCG 697

Qy    730 ACACGGCCAACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCG 789
Db    698 ACACGGCTAACTACACCTGTGTTGCCAAAAACATTGTGGCCAAAAGGAAAAGCACGACAG 757

Qy    790 CTGCTGTCATCGTCTACGTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCA 849
Db    758 CAACTGTGATTGTCTATGTGAATGGAGGCTGGTCTACCTGGACCGAGTGGTCAGCGTGCA 817

Qy    850 GCGCCAGCTGTGGGCGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTC 909
Db    818 ACAGCCGCTGTGGGAGAGGCTTCCAGAAGCGCACAAGGACCTGCACTAACCCTGCCCCAC 877

Qy    910 TCAACGGGGGCGCTTTCTGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCCTCTCTGC 969
Db    878 TCAATGGGGGGGCCTTCTGCGAGGGGCAAAATGTTCAGAAAATAGCTTGCACCACCCTGT 937

Qy    970 TTGTCTCTGTGGACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACT 1029
Db    938 GT---CCAGTGGATGGCAAATGGACGTCCTGGAGCAAGTGGTCCACTTGTGGCACAGAGT 994

Qy   1030 GCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGT 1089
Db    995 GTACCCACTGGCGCCGGAGGGAGTGCACAGCTCCGGCCCCGAAGAATGGAGGCAAGGACT 1054

Qy   1090 GCCAGGGCACTGACCTGGACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTT 1149
Db   1055 GTGAGGGACTGGTGCTGCAGTCTAAGAACTGCACTGATGGGCTCTGCATGCAGGCTGCAC 1114

Qy   1150 CTGGCCCTGAGGACGTGGCCCTCTATGTGGG---CCTCATCGCCGTGGCCGTCTGCCTGG 1206
Db   1115 CTGACTCGGATGATGTTGCTCTCTACGTGGGGATTGTCATTGCTGTGATTGTGTGCCTGG 1174

Qy   1207 TCCTGCTGCTGCTTGTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAG 1266
Db   1175 CTATTTCTGTGGTTGTGGCCCTGTTTGTCTATCGCAAGAACCACCGTGACTTTGAGTCAG 1234

Qy   1267 ATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCA 1326
Db   1235 ATATTATCGACTCATCGGCGCTAAATGGGGGATTTCAGCCTGTTAACATCAAGGCTGCAA 1294

Qy   1327 AAGCAGACAACCCCCATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCACCTACC 1386
Db   1295 GACAAGAC---------CTCTTGGCAGTGCCACCAGACCTCACTTCTGCTGCAGCCATGT 1345

Qy   1387 AGGGCAGTCTCTGTCCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGC 1446
Db   1346 ACAGGGGGCCTGTGTATGCCTTGCATGATGTCTCTGATAAAATCCCAATGACCAATTCTC 1405

Qy   1447 ACCTGCTCAGCCCCCTGGGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTG 1506
Db   1406 CGATCCTGGACCCACTGCCCAATCTGAAGATTAAAGTTTATAACACCTCTGGAGCAGTCA 1465

Qy   1507 AGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC-------------------------- 1538
Db   1466 CCCCCCAGGATGAACTCTCTGACTTCTCCTCCAAGCTGTCCCCACAGATTACCCAGTCTC 1525

Qy   1539 -------------------------CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCA 1572
Db   1526 TGTTGGAGAATGAGACTCTGAACGTGAAGAACCAAAGCCTTGCACGGCAAACAGACCCAT 1585

Qy   1573 GCAACATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAG 1632
Db   1586 CCTGCACTGCATTTGGGACCTTCAACTCGTTAGGGGGCCACCTAGTAATTCCTAATTCAG 1645

Qy   1633 GTATCAGCCTCCTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACC 1692
Db   1646 GAGTGAGCTTGCTGATCCCAGCAGGGGCTGTTCCCCAAGGAAGAGTCTATGAAATGTATG 1705

Qy   1693 TCACGCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGA 1752
Db   1706 TGACAGTCCACAGGAAGGAGGGCATGAGACCACCTGTAGAAGACAGCCAGACGCTGCTGA 1765

Qy   1753 GTCCCATCGTTAGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTA 1812
Db   1766 CACCAGTGGTGAGCTGTGGCCCACCAGGAGCGCTGCTGACCCGACCCGTTGTGCTGACCA 1825

Qy   1813 TGGACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGT 1872
Db   1826 TGCACCACTGTGCTGAGCCCAACATGGATGACTGGCAGATCCAGCTCAAGCACCAGGCAG 1885

Qy   1873 GCGAGGGCAGCTGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCT 1932
Db   1886 GCCAGGGACCATGGGAGGATGTAGTGGTGGTC---GGGGAGGAAAACTTCACCACTCCAT 1942

Qy   1933 ACTACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTG 1992
Db   1943 GCTACATCCAGCTGGACCCAGAGGCCTGTCATATCCTGACGGAGACCCTCAGCACGTACG 2002

Qy   1993 CCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTG 2052
Db   2003 CCTTGGTGGGACAATCCATCACCAAAGCAGCAGCCAAACGTCTCAAATTGGCCATCTTTG 2062

Qy   2053 CGCCGGTGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCC 2112
Db   2063 GACCACTGTCCTGTTCCTCACTGGAGTACAGCATCCGCGTCTACTGCCTCGATGACACAC 2122

Qy   2113 ACGATGCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGG 2172
Db   2123 AGGATGCCCTGAAGGAGGTCCTCCAGCTTGAGCGGCAGATGGGTGGGCAGCTGTTGGAGG 2182

Qy   2173 AGCCACGGGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATG 2232
Db   2183 AACCCAAAACTTTGCATTTTAAGGGAAGTACCCACAACCTGCGCTTATCCATTCATGACA 2242

Qy   2233 TGCCCAGCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATC 2292
Db   2243 TTGCCCACTCTCTCTGGAAGAGCAAACTGCCGGCTAAATACCAGGAGATTCCTTTTTACC 2302

Qy   2293 ACATCTGGAATGGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCC 2352
Db   2303 ACATCTGGAGTGGGTGCCAGAGGAACTTGCACTGCACCTTCACGCTGGAACGATTCAGTC 2362

Qy   2353 CCAGCACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGA 2412
```

```
Db       2363 TCAATACCCTGGAGCTCGTCTGCAAACTCTGTGTGCGGCAAGTCGAAGGAGAAGGGCAGA 2422

Qy       2413 GCTTCAGCATCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGG 2472
Db       2423 TCTTCCAGCTGAACTGCTCAGTATCAGAGGAACCCACTGGCATTGATTATCCCATCATGG 2482

Qy       2473 AGAGTGAAGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCA 2532
Db       2483 ATTCAGCAGGCAGCATCACAACGATAGTTGGGCCCAACGCTTTCAGCATCCCCCTCCCAA 2542

Qy       2533 TTCGGCAGAAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGA 2592
Db       2543 TAAGGCAGAAGCTCTGCAGCAGCCTGGATGCACCCCAGACCCGGGGCCATGACTGGAGGA 2602

Qy       2593 CTCTGGCCCAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCC 2652
Db       2603 TGCTGGCCCACAAGCTGAAATTGGACAGGTACCTAAATTATTTGCTACGAAGTCGAGTC 2662

Qy       2653 CCACAGCCATGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCC 2712
Db       2663 CCACTGGGGTGATCCTGGATCTCTGGGAAGCCCAGAATTTCCCTGATGGCAACCTGAGCA 2722

Qy       2713 AGCTGGCTGCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGG 2772
Db       2723 TGCTGGCAGCAGTGCTGGAAGAAATGGGACGACATGAAACCGTTGTTTCTTTGGCAGCAG 2782

Qy       2773 AGGCTGAGTGCTGAGGCCGGCCAGGCCCGA 2802
Db       2783 AAGGAAATTACTGATGCAGCCTTGGCGTGA 2812
```

```
RESULT 15
AB118026
LOCUS       AB118026               9328 bp    mRNA    linear   ROD 19-DEC-2003
DEFINITION  Rattus norvegicus Unc5h3 mRNA, complete cds.
ACCESSION   AB118026
VERSION     AB118026.1  GI:40217509
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1
  AUTHORS   Kuramoto,T., Kuwamura,M. and Serikawa,T.
  TITLE     Rat neurological mutations cerebellar vermis defect and hobble are
            caused by mutations in the netrin-1 receptor gene Unc5h3
  JOURNAL   Mol. Brain Res. (2003) In press
REFERENCE   2  (bases 1 to 9328)
  AUTHORS   Kuramoto,T. and Serikawa,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-AUG-2003) Takashi Kuramoto, Institute of Laboratory
            Animals, Graduate School of Medicine, Kyoto University;
            Yoshidakonoe-cho, Sakyo-ku, Kyoto 606-8501, Japan
            (E-mail:tkuramot@anim.med.kyoto-u.ac.jp,
            URL:www.anim.med.kyoto-u.ac.jp, Tel:81-75-753-4494,
            Fax:81-75-753-4409)
FEATURES             Location/Qualifiers
     source          1..9328
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10116"
                     /chromosome="2"
     gene            1..9328
                     /gene="Unc5h3"
     CDS             151..2946
                     /gene="Unc5h3"
                     /codon_start=1
                     /protein_id="BAD05181.1"
                     /db_xref="GI:40217510"
                     /translation="MRKGLRATAARCGLGLGYLLQMLVLPALALLSASGTGSAAQDDD
                     FFHELPETFPSDPPEPLPHFLIEPEEAYIVKNKPVNLYCKASPATQIYFKCNSEWVHQ
                     KDHVVDERVDETSGLIVREVSIEISRQQVEELFGPEDYWCQCVAWSSAGTTKSRKAYV
                     RIAYLRKTFEQEPLGKEVSLEQEVLLQCRPPEGIPMAEVEWLKNEDIIDPVEDRNFYI
                     TIDHNLIIKQARLSDTANYTCVAKNIVAKRKSTTATVIVYVNGGWSTWAEWSVCNSRC
                     GRGYQKRTRTCTNPAPLNGGAFCEGQSVQKIACTTLCPVDGRWTSWSKWSTCGTECTH
                     WRRRECTAPAPKNGGKDCDGLVLQSKNCTDGLCMQAAPDSDDVALYVGIVIAVTVCLA
                     ITVVVALFVYRKNHRDFESNIIDSSALNGGFQPVNIKAARQDLLAVPPDLTSAAAMYR
                     GPVYALHDVSDKIPMTNSPILDPLPNLKIKVYNSSGAVTPQDDLAEFSSKLSPQMTQS
                     LLENEALNLKNQSLARQTDPSCTAFGTFNSLGGHLIIPNSGVSLLIPAGAIPQGRVYE
                     MYVTVHRKENMRPPMEDSQTLLTPVVSCGPPGALLTRPVILTLHHCADPNTEDWKIQL
                     KNQAVQGQWEDVVVVGEENFTTPCYIQLDAEACHILTENLSTYALVGQSTTKAAAKRL
                     KLAIFGPLCCSSLEYSIRVYCLDDTQDALKEVLQLERQMGGQLLEEPKALHFKGSIHN
                     LRLSIHDITHSLWKSKLLAKYQEIPFYHIWSGSQRNLHCTFTLERLSLNTVELVCKLC
                     VRQVEGEGQIFQLNCTVSEEPTGIDLPLLDPASTITTVTGPSAFSIPLPIRQKLCSSL
                     DAPQTRGHDWRMLAHKLNLDRYLNYFATKSSPTGVILDLWEAQNFPDGNLSMLAAVLE
                     EMGRHETVVSLAAEGQY"
     3'UTR           2947..9328
                     /gene="Unc5h3"
     repeat_region   8727..8772
                     /note="simple sequence repeats"
                     /rpt_unit="gt"
     polyA_signal    9309..9314
                     /gene="Unc5h3"
ORIGIN

  Query Match             32.5%;  Score 936.8;  DB 10;  Length 9328;
  Best Local Similarity   61.8%;  Pred. No. 3.6e-145;
  Matches 1631;  Conservative    0;  Mismatches  937;  Indels   72;  Gaps    6;

Qy        204 CCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAG 263
Db        322 CCTGAGCCATTACCACACTTTCTCATTGAGCCCGAGGAAGCTTACATTGTGAAGAACAAG 381

Qy        264 CCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGG 323
Db        382 CCTGTGAACCTGTATTGTAAAGCCAGCCCTGCCACCCAGATCTACTTCAAGTGCAACAGC 441

Qy        324 GAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAG 383
Db        442 GAATGGGTTCATCAGAAGGACCACGTAGTAGACGAGAGGGTAGATGAAACCTCCGGTCTC 501

Qy        384 CCGACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTG 443
Db        502 ATTGTGAGAGAAGTGAGCATTGAGATTTCACGCCAGCAGGTGGAGGAACTGTTCGGGCCT 561

Qy        444 GAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAG 503
Db        562 GAAGATTACTGGTGCCAGTGTGTGGCCTGGAGCTCAGCGGGCACTACGAAGAGTCGGAAG 621

Qy        504 GCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAG 563
Db        622 GCATACGTGCGCATTGCGTATCTGCGGAAGACATTTGAGCAGGAACCCTTGGGAAAAGAA 681

Qy        564 GTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCC 623
Db        682 GTGTCCTTGGAGCAGGAAGTGTTACTCCAGTGCCGGCCACCTGAAGGGATCCCAATGGCT 741

Qy        624 GAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATAC 683
Db        742 GAGGTGGAGTGGCTAAAGAATGAAGACATAATTGATCCTGTCGAAGATCGGAATTTTTAT 801

Qy        684 ATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTAC 743
Db        802 ATCACTATCGATCACAACCTGATCATCAAGCAGGCACGACTCTCAGATACAGCGAATTAT 861

Qy        744 ACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTC 803
Db        862 ACCTGTGTTGCCAAAAACATTGTTGCCAAGAGGAAAAGCACCACAGCCACCGTCATCGTG 921

Qy        804 TACGTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGG 863
Db        922 TATGTTAATGGTGGCTGGTCCACCTGGGCAGAATGGTCTGTGTGTAACAGCCGCTGTGGG 981

Qy        864 CGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCT 923
```

```
Db    982 CGAGGATATCAGAAACGCACAAGAACCTGCACCAACCCAGCCCCCCTCAATGGTGGGGCC 1041

Qy    924 TTCTGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGAC 983
Db   1042 TTCTGCGAGGGGCAGAGTGTACAGAA---AATAGCATGCACCACGTTATGCCCAGTGGAT 1098

Qy    984 GGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGG 1043
Db   1099 GGCAGGTGGACTTCGTGGAGCAAATGGTCAACTTGTGGGACTGAGTGCACCCACTGGCGG 1158

Qy   1044 AGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGAC 1103
Db   1159 AGGAGGGAGTGTACAGCACCAGCCCCCAAGAACGGGGGTAAGGACTGTGATGGCCTGGTC 1218

Qy   1104 CTGGACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGAC 1163
Db   1219 CTCCAATCCAAGAACTGCACCGATGGGCTGTGCATGCAGGCTGCTCCTGACTCAGATGAC 1278

Qy   1164 GTGGCCCTCTATGTGGG---CCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTT 1220
Db   1279 GTTGCTCTCTACGTTGGGATTGTGATCGCTGTAACCGTTTGTCTGGCAATCACTGTTGTG 1338

Qy   1221 GTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCG 1280
Db   1339 GTGGCCCTGTTCGTGTATCGCAAGAACCACCGTGACTTTGAGTCTAATATCATTGACTCC 1398

Qy   1281 TCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCC 1340
Db   1399 TCAGCACTCAATGGCGGCTTTCAGCCTGTGAACATCAAG---------GCTGCCAGACAA 1449

Qy   1341 CATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGT 1400
Db   1450 GATCTCCTGGCAGTACCCCCGGATCTCACCTCAGCTGCAGCCATGTACAGGGGACCCGTC 1509

Qy   1401 CCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCC 1460
Db   1510 TATGCTCTCCATGATGTCTCAGACAAAATCCCAATGACCAATTCTCCAATTCTAGACCCA 1569

Qy   1461 CTGGGTGGCGGCCGCCACACACTGCACCACAGCT------------CTCCCACCTCTGAG 1508
Db   1570 CTGCCCAATTTGAAAATCAAAGTGTACAACAGCTCAGGTGCTGTCACTCCTCAGGACGAC 1629

Qy   1509 GCCGAGGAGTTCGTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGC 1568
Db   1630 CTTGCTGAGTTCTCATCCAAGCTGTCACCCCAGATGACCCAATCCTTGCTAGAGAATGAG 1689

Qy   1569 ACCAGCAACATGA----------------------------------------CCTAT 1586
Db   1690 GCCCTGAACCTGAAGAACCAGAGTCTCGCAAGACAGACTGACCCATCCTGCACAGCGTTT 1749

Qy   1587 GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTC 1646
Db   1750 GGCACCTTCAACTCCCTGGGCGGTCACCTCATCATCCCTAATTCAGGAGTAAGCTTGCTG 1809

Qy   1647 ATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG 1706
Db   1810 ATTCCCGCTGGGGCCATTCCTCAGGGGAGAGTCTATGAAATGTATGTGACTGTACACAGG 1869

Qy   1707 CCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC 1766
Db   1870 AAAGAAAATATGAGGCCCCCCATGGAAGACTCTCAGACCCTGCTGACCCCTGTGGTGAGC 1929

Qy   1767 TGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG 1826
Db   1930 TGTGGGCCTCCTGGAGCTCTGTTGACCCGTCCTGTCATCCTCACTCTGCATCACTGTGCA 1989

Qy   1827 GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG 1886
Db   1990 GACCCTAACACCGAGGACTGGAAGATCCAGCTCAAAAACCAGGCAGTGCAGGGACAGTGG 2049

Qy   1887 GAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTG 1946
Db   2050 GAGGATGTTGTGGTTG---TTGGAGAGGAAAACTTCACAACCCCCTGTTACATTCAGCTG 2106

Qy   1947 GAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAG 2006
Db   2107 GATGCAGAGGCTTGCCATATTCTCACGGAGAACCTCAGTACCTATGCCCTGGTTGGGCAG 2166

Qy   2007 GCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGC 2066
Db   2167 TCCACCACCAAAGCAGCTGCCAAGCGTCTGAAGCTGGCCATCTTTGGGCCCCTCTGCTGC 2226

Qy   2067 ACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAG 2126
Db   2227 TCTTCCCTGGAGTATAGCATCAGAGTCTACTGCCTGGATGACACACAGGATGCCCTGAAG 2286

Qy   2127 GAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTG 2186
Db   2287 GAAGTTCTACAACTGGAAAGGCAAATGGGAGGACAGCTCCTAGAAGAACCCAAGGCTCTT 2346

Qy   2187 CACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTG 2246
Db   2347 CATTTTAAAGGCAGCATCCACAACCTGCGCCTCTCAATTCATGACATCACCCATTCCCTC 2406

Qy   2247 TGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGC 2306
Db   2407 TGGAAGAGCAAATTGCTGGCTAAGTATCAGGAAATTCCATTTTACCACATCTGGAGTGGC 2466

Qy   2307 ACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGAC 2366
Db   2467 TCTCAAAGAAACCTCCACTGTACCTTCACTCTGGAAAGACTCAGCCTAAACACAGTAGAG 2526

Qy   2367 CTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAAC 2426
Db   2527 CTGGTTTGCAAACTCTGTGTGCGGCAGGTTGAAGGAGAAGGGCAGATCTTCCAGCTCAAC 2586

Qy   2427 TTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGG 2486
Db   2587 TGTACTGTGTCAGAGGAACCTACTGGTATCGATTTGCCTCTCCTGGACCCTGCTAGCACC 2646

Qy   2487 GTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATA 2546
Db   2647 ATTACCACTGTCACTGGACCAAGTGCTTTCAGCATCCCTCTTCCTATCCGGCAGAAGCTA 2706

Qy   2547 ATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAA 2606
Db   2707 TGTAGCAGCCTGGATGCCCCCAAACAAGAGGCCATGACTGGAGGATGCTGGCCCACAAA 2766

Qy   2607 CTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATC 2666
Db   2767 CTTAACCTGGACAGGTACTTGAATTACTTTGCCACCAAATCGAGCCCGACTGGCGTAATC 2826

Qy   2667 CTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCA 2726
Db   2827 CTGGATCTCTGGGAAGCACAGAACTTCCCCGATGGAAACCTGAGCATGCTGGCAGCCGTC 2886

Qy   2727 GTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2786
Db   2887 CTGGAAGAAATGGAAGACACGAAACCGTAGTGTCCTTGGCAGCAGAAGGTCAGTATTGA 2946
```

Search completed: October 6, 2004, 15:50:36
Job time : 11154 secs

```
                 GenCore version 5.1.6
        Copyright (c) 1993 - 2004  Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:         October  6, 2004, 09:36:51 ; Search time 1081 Seconds
                                             (without alignments)
                                             11321.978 Million cell updates/sec

Title:          US-09-970-944-1
Perfect score:  2881
Sequence:       1 agctggggctccgggctgag..........ccttccccacaccggggaga 2881

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters:      6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       N_Geneseq_29Jan04:*
                 1:  geneseqn1980s:*
                 2:  geneseqn1990s:*
                 3:  geneseqn2000s:*
                 4:  geneseqn2001as:*
                 5:  geneseqn2001bs:*
                 6:  geneseqn2002s:*
                 7:  geneseqn2003as:*
                 8:  geneseqn2003bs:*
                 9:  geneseqn2003cs:*
                 10:  geneseqn2004s:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                              SUMMARIES
                     %
Result             Query
  No.     Score    Match Length DB  ID                           Description
-------------------------------------------------------------------------------
   1      2881    100.0    2881  6  ABK49422                 Abk49422 DNA encod
   2    2676.4     92.9    2752  6  ABK37922                 Abk37922 cDNA enco
   3      2615     90.8    2697  6  ABK52891                 Abk52891 Human net
   4    2402.6     83.4    3580  6  ABK15169                 Abk15169 Human REP
   5    2296.2     79.7    2907  4  AAK52261                 Aak52261 Human pol
   6    2235.8     77.6    3014  2  AAV52940                 Aav52940 Rat UNC-5
   7    2189.4     76.0    2697  6  AAS16843                 Aas16843 Rat netri
   8    1570.4     54.5    1787  2  AAV52941                 Aav52941 Human UNC
   9    1291.2     44.8    9700  4  AAC90958                 Aac90958 Plasmid p
  10    1193.6     41.4    1321  4  AAH99530                 Aah99530 Human pro
  11       916     31.8    3646  5  AAS75738                 Aas75738 DNA encod
  12       889     30.9    2860  6  ABT06279                 Abt06279 Human NOV
  13     887.4     30.8    2860  6  ABT06280                 Abt06280 Human NOV
  14     875.8     30.4    2995  6  ABK92062                 Abk92062 DNA encod
  15     866.4     30.1    2895  6  ABQ93898                 Abq93898 Human tra
  16     853.6     29.6    3884  4  AAS21316                 Aas21316 Human cDN
  17     853.6     29.6    3884  7  ACD23925                 Acd23925 Novel hum
  18     853.6     29.6    3884  7  ACA67066                 Aca67066 cDNA enco
  19     853.6     29.6    3884  7  ACA03675                 Aca03675 cDNA enco
  20     853.6     29.6    3884  7  ABX89213                 Abx89213 DNA encod
  21     853.6     29.6    3884  7  ACD41867                 Acd41867 Human sec
  22     853.6     29.6    3884  7  ACA04096                 Aca04096 Human cDN
  23     853.6     29.6    3884  8  ADA45664                 Ada45664 Novel hum
  24     853.6     29.6    3884  8  ADA76095                 Ada76095 Human PRO
  25     853.6     29.6    3884  8  ADA18745                 Ada18745 Human PRO
  26     853.6     29.6    3884  8  ADA61368                 Ada61368 Homo sapi
  27     853.6     29.6    3884  8  ADB19153                 Adb19153 Novel hum
  28     853.6     29.6    3884  8  ADB27694                 Adb27694 cDNA enco
  29     853.6     29.6    3884  8  ADA86173                 Ada86173 Novel hum
  30     853.6     29.6    3884  8  ADB15737                 Adb15737 Human PRO
  31     853.6     29.6    3884  8  ADA47523                 Ada47523 Human PRO
  32     853.6     29.6    3884  8  ADA67318                 Ada67318 Human PRO
  33     853.6     29.6    3884  8  ADB30325                 Adb30325 cDNA enco
  34     853.6     29.6    3884  8  ADA85621                 Ada85621 Novel hum
  35     853.6     29.6    3884  8  ADA96833                 Ada96833 Human PRO
  36     853.6     29.6    3884  8  ADA79137                 Ada79137 Human PRO
  37     853.6     29.6    3884  8  ADA87276                 Ada87276 Novel hum
  38     853.6     29.6    3884  8  ADB16478                 Adb16478 Human PRO
  39     853.6     29.6    3884  8  ADA91570                 Ada91570 Novel hum
  40     853.6     29.6    3884  8  ADB14633                 Adb14633 Human PRO
  41     853.6     29.6    3884  8  ADB18594                 Adb18594 Novel hum
  42     853.6     29.6    3884  8  ADA93809                 Ada93809 Human PRO
  43     853.6     29.6    3884  8  ADB19705                 Adb19705 Novel hum
  44     853.6     29.6    3884  8  ADB13017                 Adb13017 Human PRO
  45     853.6     29.6    3884  8  ACD98496                 Acd98496 Novel hum

                              ALIGNMENTS

RESULT 1
ABK49422
ID   ABK49422 standard; DNA; 2881 BP.
XX
AC   ABK49422;
XX
DT   15-JUL-2002  (first entry)
XX
DE   DNA encoding human UNC5-like protein NOV1.
XX
KW   Human; NOVX polypeptide; cardiomyopathy; atherosclerosis; cancer;
KW   cell signal processing; metabolic pathway modulation; cancerous tissue;
KW   antibody; diabetes; transgenic animal; UNC5-like protein; NOV1;
KW   chromosome 13; gene; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             87. .2786
FT                   /*tag=  a
FT                   /product= "Human UNC5-like protein NOV1"
XX
PN   WO200229038-A2.
XX
PD   11-APR-2002.
XX
PF   04-OCT-2001; 2001WO-US031377.
XX
PR   04-OCT-2000; 2000US-0237862P.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Herrmann JL,  Rastelli L,  Shimkets RA;
XX
DR   WPI; 2002-340104/37.
DR   P-PSDB; AAU79939.
XX
PT   Novel isolated NOVX polypeptide, and encoded polynucleotide, useful for
PT   treating cardiomyopathy, artherosclerosis, and cancer.
XX
PS   Claim 8; Page 7-8; 180pp; English.
XX
CC   The present invention relates to a new NOVX polypeptide having a 900
CC   (NOV1), 4349 (NOV2), 940 (NOV3), 798 (NOV4), 865 (NOV5), or 331 (NOV6)
CC   residue amino acid sequence, as given in the specification. The novel
CC   polypeptide, and its encoding polynucleotide, are used to treat
```

```
CC   cardiomyopathy, atherosclerosis, cancer or a disease related to cell
CC   signal processing and metabolic pathway modulation, in a human. Detecting
CC   the polypeptide or polynucleotide is useful for identifying cancerous
CC   tissue. The antibody can be used to treat diabetes or cancer. The host
CC   cells can be used to produce non-human transgenic animals useful in drug
CC   screening. The present nucleic acid sequence is that of the human UNC5-
CC   like NOV1 gene located on chromosome 13. This sequence encodes the human
CC   UNC5-like protein NOV1 of the invention
XX
SQ   Sequence 2881 BP; 526 A; 985 C; 868 G; 502 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 2881;  DB 6;  Length 2881;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 2881;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGCTGGGGCTCCGGGCTGAGGCGCTAAAGCCGCCCTCCCGCCCGCGGGGCCCCGCGCCCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGCTGGGGCTCCGGGCTGAGGCGCTAAAGCCGCCCTCCCGCCCGCGGGGCCCCGCGCCCG 60

Qy         61 GCCCGCCCGCCTGCCCGCCCGCGGCCATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCCCGCCCGCCTGCCCGCCCGCGGCCATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCC 120

Qy        121 TGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGG 180

Qy        181 CCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGG 240

Qy        241 ATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGC 300

Qy        301 AGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCA 360

Qy        361 GCACAGACGGGAGCAGTGGTGAGCCGACCATGGAGGTCCGCATTAATGTCTCAAGGCAGC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCACAGACGGGAGCAGTGGTGAGCCGACCATGGAGGTCCGCATTAATGTCTCAAGGCAGC 420

Qy        421 AGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCT 480

Qy        481 CGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCG 540

Qy        541 AGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTC 600

Qy        601 CACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACC 660

Qy        661 CGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCC 720

Qy        721 GCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCA 780

Qy        781 GCGCCTCCGCTGCTGTCATCGTCTACGTGAACGGTGGGTGGTCGACGTGGACCGAGTGGT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GCGCCTCCGCTGCTGTCATCGTCTACGTGAACGGTGGGTGGTCGACGTGGACCGAGTGGT 840

Qy        841 CCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAACC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAACC 900

Qy        901 CGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAATGTCCATGACCGCACCGTCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAATGTCCATGACCGCACCGTCT 960

Qy        961 CCTCTCTGCTTGTCTCTGTGGACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCTCTCTGCTTGTCTCTGTGGACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTG 1020

Qy       1021 GGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAG 1080

Qy       1081 GGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGTACCAGTGACCTCTGTGTAC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGTACCAGTGACCTCTGTGTAC 1140

Qy       1141 ACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCT 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 ACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCT 1200

Qy       1201 GCCTGGTCCTGCTGCTGCTTGTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GCCTGGTCCTGCTGCTGCTTGTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGG 1260

Qy       1261 ACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 ACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGC 1320

Qy       1321 CCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCA 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCA 1380

Qy       1381 CCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCA 1440

Qy       1441 ATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 ATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCA 1500

Qy       1501 CCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGC 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGC 1560

Qy       1561 CCCGAGGCACCAGCAACATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGA 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CCCGAGGCACCAGCAACATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGA 1620

Qy       1621 TCCCTAATACAGGTATCAGCCTCCTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCT 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 TCCCTAATACAGGTATCAGCCTCCTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCT 1680

Qy       1681 ATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTC 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 ATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTC 1740

Qy       1741 AGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAG 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 AGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAG 1800

Qy       1801 TCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCA 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 TCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCA 1860

Qy       1861 AAAAGCAGTCGTGCGAGGGCAGCTGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGC 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 AAAAGCAGTCGTGCGAGGGCAGCTGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGC 1920

Qy       1921 CCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGC 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 CCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGC 1980
```

```
Qy       1981 TGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGC 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 TGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGC 2040

Qy       2041 TGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCC 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 TGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCC 2100

Qy       2101 TGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGAC 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 TGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGAC 2160

Qy       2161 AGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTAT 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 AGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTAT 2220

Qy       2221 CCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGA 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 CCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGA 2280

Qy       2281 TCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGG 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 TCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGG 2340

Qy       2341 AGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGG 2400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 AGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGG 2400

Qy       2401 GCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGC 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 GCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGC 2460

Qy       2461 TGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGA 2520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 TGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGA 2520

Qy       2521 TCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTG 2580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 TCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTG 2580

Qy       2581 CCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCT 2640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 CCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCT 2640

Qy       2641 CCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACG 2700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 CCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACG 2700

Qy       2701 GCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCT 2760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 GCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCT 2760

Qy       2761 TCACAGTGTCGGAGGCTGAGTGCTGAGGCCGGCCAGGCCCGACACCTACACTCTCACCAG 2820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 TCACAGTGTCGGAGGCTGAGTGCTGAGGCCGGCCAGGCCCGACACCTACACTCTCACCAG 2820

Qy       2821 CTTTGGCACCCACCAAGGACAGGCAGAAGCCGGACAGGGGCCCTTCCCCACACCGGGGAG 2880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 CTTTGGCACCCACCAAGGACAGGCAGAAGCCGGACAGGGGCCCTTCCCCACACCGGGGAG 2880

Qy       2881 A 2881
              |
Db       2881 A 2881

RESULT 2
ABK37922
ID   ABK37922 standard; cDNA; 2752 BP.
XX
AC   ABK37922;
XX
DT   21-MAY-2002  (first entry)
XX
DE   cDNA encoding Human protein NOV1.
XX
KW   Human; NOVX; ss; gene; cardiomyopathy; atherosclerosis; diabetes;
KW   cell signal processing disorder; metabolic disorder; obesity; infection;
KW   anorexia; cancer-associated cachexia; cancer; neurodegenerative disorder;
KW   Alzheimer's disease; Parkinson's disease; immune disorder;
KW   haematopoietic disorders; dyslipidaemia; pain; asthma; hypertension;
KW   osteoporosis; Crohn's disease; multiple sclerosis; angina pectoris;
KW   myocardial infarction; ulcer; allergy; benign prostatic hypertrophy;
KW   psychosis; neurological disorder; anxiety; schizophrenia;
KW   manic depression; dementia; dyskinesia; Huntington's disease;
KW   Gilles de la Tourette's syndrome; gene therapy.
XX
OS   Homo sapiens.
XX
PN   WO200210216-A2.
XX
PD   07-FEB-2002.
XX
PF   30-JUL-2001; 2001WO-US024225.
XX
PR   28-JUL-2000; 2000US-0221409P.
PR   04-AUG-2000; 2000US-0222840P.
PR   04-AUG-2000; 2000US-0223752P.
PR   04-AUG-2000; 2000US-0223762P.
PR   04-AUG-2000; 2000US-0223769P.
PR   04-AUG-2000; 2000US-0223770P.
PR   14-AUG-2000; 2000US-0225146P.
PR   15-AUG-2000; 2000US-0225392P.
PR   15-AUG-2000; 2000US-0225470P.
PR   16-AUG-2000; 2000US-0225697P.
PR   01-FEB-2001; 2001US-0263662P.
PR   05-APR-2001; 2001US-0281645P.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Padigaru M,  Mezes P,  Mishra V,  Burgess C,  Casman S,  Grosse WM;
PI   Alsobrook JP,  Lepley DM,  Gerlach VL,  Macdougall JR,  Smithson G;
XX
DR   WPI; 2002-180074/23.
DR   P-PSDB; AAU85403.
XX
PT   New isolated cytoplasmic, nuclear, membrane bound, or secreted
PT   polypeptide, useful for treating cardiomyopathy, atherosclerosis,
PT   infections, cancer, neurodegenerative, metabolic, hematopoietic and
PT   immune disorders.
XX
PS   Claim 9; Page 9-10; 213pp; English.
XX
CC   The invention relates to an isolated cytoplasmic, nuclear, membrane
CC   bound, or secreted polypeptide (NOVX, x= 1-14) their variants or mature
CC   form. Also included are the nucleic acids encoding the NOVX proteins, a
CC   vector comprising the nucleic acid, a cell comprising the vector, an anti
CC   -NOVX antibody and modulators of NOVX. NOVX, the nucleic acid and the
CC   antibody are useful for treating or preventing a NOVX-associated
CC   disorder, where the disorder is selected from cardiomyopathy,
CC   atherosclerosis, diabetes, a disorder related to cell signal processing
CC   and metabolic pathway modulation, metabolic disorders, obesity,
CC   infectious disease, anorexia, cancer-associated cachexia, cancer,
CC   neurodegenerative disorders, Alzheimer's disease, Parkinson's disease,
CC   immune disorders, haematopoietic disorders, and the various
CC   dyslipidaemias, metabolic disturbances associated with obesity, the
CC   metabolic syndrome X and wasting disorders associated with chronic
CC   diseases, bacterial, fungal, protozoal and viral infections, pain,
CC   bulimia, asthma, hypertension, urinary retention, osteoporosis, Crohn's
CC   disease, multiple sclerosis, Albright Hereditary Osteodystrophy, angina
CC   pectoris, myocardial infarction, ulcer, allergy, benign prostatic
CC   hypertrophy, and psychotic and neurological disorders, including anxiety,
CC   schizophrenia, manic depression, delirium, dementia, and dyskinesias,
CC   such as Huntington's disease and Gilles de la Tourette's syndrome. The
CC   nucleic acid is useful in gene therapy. The present sequence encodes a
CC   NOVX protein
```

```
XX
SQ   Sequence 2752 BP; 505 A; 937 C; 829 G; 481 T; 0 U; 0 Other;

  Query Match             92.9%;  Score 2676.4;  DB 6;  Length 2752;
  Best Local Similarity   98.9%;  Pred. No. 0;
  Matches 2728;  Conservative    0;  Mismatches   21;  Indels    9;  Gaps    3;

Qy         42 CCGCGGGGCCCCGCGCCCGGCCCGCCCGCCTGCCCGCCCGCGGCCATGGCCGTCCGGCCC 101
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCGCGGGGCCCCGCGCCCGGCCCGCCCGCCTGCCCGCCCGCGGCCATGGCCGTCCGGCCC 60

Qy        102 GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC 161
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC 120

Qy        162 CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC 221
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC 180

Qy        222 TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC 281
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC 240

Qy        282 AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG 341
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG 300

Qy        342 GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCGACCATGGAGGTCCGC 401
              ||||||||||||||||||||||||||||||||||||||   ||| |||||||||||||||
Db        301 GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCCACCATGGAGGTCCGC 360

Qy        402 ATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAG 461
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAG 420

Qy        462 TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC 521
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC 480

Qy        522 AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC 581
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC 540

Qy        582 ATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG 641
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG 600

Qy        642 AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC 701
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC 660

Qy        702 CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC 761
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC 720

Qy        762 ATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTACGTGAACGGTGGGTGG 821
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTACGTGAACGGTGGGTGG 780

Qy        822 TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG 881
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG 840

Qy        882 AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT 941
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT 900

Qy        942 GTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGCAGCTGGAGCCCGTGG 1001
              |||||  |    || | || | |    ||   | | || |||||||||||||||||||||
Db        901 GTCCAGAA---AACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGCTGGAGCCCGTGG 957

Qy       1002 AGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC 1061
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        958 AGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC 1017

Qy       1062 CCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT 1121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1018 CCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT 1077

Qy       1122 ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC 1181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1078 ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC 1137

Qy       1182 CTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCATCCTCGTTTATTGC 1241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1138 CTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCATCCTCGTTTATTGC 1197

Qy       1242 CGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC 1301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1198 CGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC 1257

Qy       1302 CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG 1361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1258 CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG 1317

Qy       1362 GACCTCAG---CACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC 1418
              ||||||||   |||||||||||||||||||||||||||||||||||||||||||||||||
Db       1318 GACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC 1377

Qy       1419 AGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC 1478
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1378 AGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC 1437

Qy       1479 ACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC 1538
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1438 ACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC 1497

Qy       1539 CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC 1598
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1498 CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC 1557

Qy       1599 TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCCAGAT 1658
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1558 TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCCAGAT 1617

Qy       1659 GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG 1718
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1618 GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG 1677

Qy       1719 AGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT 1778
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1678 AGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT 1737

Qy       1779 GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT 1838
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1738 GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT 1797

Qy       1839 GACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGCAGGATGTG 1898
              |||||||||||||||||||||||||||||||||||||||||||||||||   ||||||||
Db       1798 GACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGG---AGGATGTG 1854

Qy       1899 CTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC 1958
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1855 CTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC 1914

Qy       1959 TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG 2018
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1915 TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG 1974

Qy       2019 GCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAG 2078
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1975 GCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAG 2034

Qy       2079 TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG 2138
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2035 TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG 2094
```

```
Qy       2139 CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC 2198
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2095 CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC 2154

Qy       2199 AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAG 2258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2155 AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAG 2214

Qy       2259 CTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC 2318
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2215 CTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC 2274

Qy       2319 TTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAG 2378
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2275 TTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAG 2334

Qy       2379 CTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACC 2438
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2335 CTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACC 2394

Qy       2439 AAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG 2498
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2395 AAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG 2454

Qy       2499 GTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTG 2558
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2455 GTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTG 2514

Qy       2559 GACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGAC 2618
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2515 GACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGAC 2574

Qy       2619 AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGG 2678
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2575 AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGG 2634

Qy       2679 GAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTG 2738
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2635 GAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTG 2694

Qy       2739 GGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGAGGCCGGCCAG 2796
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2695 GGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGAGGCCGGCCAG 2752
```

RESULT 3
ABK52891
ID   ABK52891 standard; DNA; 2697 BP.
XX
AC   ABK52891;
XX
DT   27-AUG-2002  (first entry)
XX
DE   Human netrin binding membrane receptor UNC5H-1 DNA sequence #1.
XX
KW   Netrin binding membrane receptor; receptor; UNC5H-1; gene; ds; human;
KW   nootropic; neuroprotective; cytostatic; antiparkinsonian;
KW   cerebroprotective; cancer; central nervous system; CNS; stroke;
KW   Parkinson's disease; multiple sclerosis; Alzheimer's disease.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .2697
FT                   /*tag=  a
FT                   /product= "Netrin binding membrane receptor UNC5H-1"
XX
PN   WO200233080-A2.
XX
PD   25-APR-2002.
XX
PF   15-OCT-2001; 2001WO-EP011891.
XX
PR   16-OCT-2000; 2000US-0240061P.
XX
PA   (FARB ) BAYER AG.
XX
PI   Koehler RH;
XX
DR   WPI; 2002-463314/49.
DR   P-PSDB; AAU97899.
XX
PT   Novel human netrin binding membrane receptor polypeptide and
PT   polynucleotides for identifying modulating agents useful in treating
PT   diseases e.g. Parkinson's disease, multiple sclerosis, stroke,
PT   Alzheimer's disease.
XX
PS   Claim 1; Fig 1; 94pp; English.
XX
CC   This invention relates to the DNA and protein sequences of a novel
CC   purified human netrin binding membrane receptor, UNC5H-1. The DNA
CC   sequence of the invention is useful as a probe for detecting a nucleic
CC   acid encoding the UNC5H-1 protein in a biological sample. The sequences
CC   of the invention are useful to screen for agents which decrease the
CC   activity of the UNC5H-1 protein. The sequences are also useful for
CC   screening agents which regulate (modulate) the activity of the protein of
CC   the invention. A pharmaceutical composition containing the protein of the
CC   invention or a reagent that modulates the activity of the UNC5H-1 protein
CC   may be useful for treating a UNC5H-1 dysfunction related disease such as
CC   cancer or a central nervous system (CNS) disorders (e.g, Parkinson's
CC   disease, multiple sclerosis, stroke and Alzheimer's disease). Fusion
CC   proteins comprising the UNC5H-1 protein are useful for generating
CC   antibodies and for in various assay systems, and the protein can be used
CC   as a bait protein in a two-hybrid assay or three-hybrid assay. The method
CC   of the invention is useful for detecting a coding sequence for the UNC5H-
CC   1 protein. The present sequence represents a DNA sequence encoding the
CC   human netrin binding membrane receptor UNC5H-1 protein of the invention
XX
SQ   Sequence 2697 BP; 503 A; 906 C; 807 G; 481 T; 0 U; 0 Other;

  Query Match             90.8%;  Score 2615;  DB 6;  Length 2697;
  Best Local Similarity   98.7%;  Pred. No. 0;
  Matches 2669;  Conservative    0;  Mismatches   25;  Indels    9;  Gaps    3;

```
Qy         87 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 146
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 60

Qy        147 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 206
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120

Qy        207 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 266
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 180

Qy        267 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 326
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 240

Qy        327 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCG 386
              ||||||||||||||||||||||||||||||||||||||||||||||||||    |||
Db        241 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 300

Qy        387 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 446
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 360

Qy        447 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 506
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 420

Qy        507 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 566
              |||||||||||||||    |||||||||||||||||||||||||||||||||||||||||
Db        421 TACATCCGCATAGCCTATTTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 480
```

```
Qy    567 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 626
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 540

Qy    627 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 686
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 600

Qy    687 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 746
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 660

Qy    747 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 806
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 720

Qy    807 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 866
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 780

Qy    867 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 926
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 840

Qy    927 TGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGC 986
          ||||||||||||||||||||  |    || | || | |    ||   | | |||||||||
Db    841 TGTGAGGGGCAGAATGTCCAGAA---AACAGCCTGCGCCACCCTGTGCCCAGTGGACGGC 897

Qy    987 AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC 1046
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    898 AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC 957

Qy   1047 CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG 1106
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    958 CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG 1017

Qy   1107 GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG 1166
          ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db   1018 GACACCCGCAACTGTACCAGTGACCTCTGTGTACACACTGCTTCTGGCCCTGAGGACGTG 1077

Qy   1167 GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC 1226
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1078 GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC 1137

Qy   1227 ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1286
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1138 ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1197

Qy   1287 CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1346
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1198 CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1257

Qy   1347 CTCACCATCCAGCCGGACCTCAG---CACCACCACCACCTACCAGGGCAGTCTCTGTCCC 1403
          |||||||||||||||||||||||   ||||||||||||||||||||||||||||||||||
Db   1258 CTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCC 1317

Qy   1404 CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG 1463
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1318 CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG 1377

Qy   1464 GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC 1523
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1378 GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC 1437

Qy   1524 TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC 1583
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1438 TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC 1497

Qy   1584 TATGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC 1643
          ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db   1498 TATGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGAATCAGCCTC 1557

Qy   1644 CTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC 1703
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1558 CTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC 1617

Qy   1704 AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT 1763
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1618 AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT 1677

Qy   1764 AGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT 1823
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1678 AGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT 1737

Qy   1824 GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC 1883
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1738 GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC 1797

Qy   1884 TGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG 1943
          ||||   |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1798 TGGG---AGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG 1854

Qy   1944 CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA 2003
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1855 CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA 1914

Qy   2004 GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC 2063
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1915 GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC 1974

Qy   2064 TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC 2123
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1975 TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC 2034

Qy   2124 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTC 2183
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2035 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTC 2094

Qy   2184 CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCC 2243
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2095 CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCC 2154

Qy   2244 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2303
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2155 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2214

Qy   2304 GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2363
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2215 GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2274

Qy   2364 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2423
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2275 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2334

Qy   2424 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2483
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2335 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2394

Qy   2484 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2543
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2395 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2454

Qy   2544 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2603
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2455 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2514

Qy   2604 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG 2663
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2515 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG 2574

Qy   2664 ATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2723
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2575 ATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2634

Qy   2724 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2783
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2635 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2694

Qy        2784 TGA 2786
               |||
Db        2695 TGA 2697

RESULT 4
ABK15169
ID   ABK15169 standard; DNA; 3580 BP.
XX
AC   ABK15169;
XX
DT   23-APR-2002  (first entry)
XX
DE   Human REPTR 1 cDNA sequence.
XX
KW   REPTR; human; antiinflammatory; cytostatic; immunosuppressive; antiviral;
KW   anti-HIV; antiarthritic; anticonvulsant; nootropic; neuroprotective;
KW   antiallergic; antibody; immunogen; endometriosis;
KW   gastrointestinal disorder; gastritis; oesophageal carcinoma;
KW   Crohn's disease; irritable bowel syndrome; ulcerative colitis;
KW   endocrine disorder; hypothalamus disorder; Kallman's disease;
KW   autoimmune disease; inflammatory disease; infertility; receptor;
KW   acquired immune deficiency syndrome; AIDS; rheumatoid arthritis; allergy;
KW   osteoarthritis; diabetes mellitus; multiple sclerosis;
KW   systemic lupus erythematosus; cell proliferative disorder; cancer;
KW   developmental disorder; Duchenne muscular dystrophy; gene;
KW   Becker muscular dystrophy; neurological disorder; epilepsy;
KW   Alzheimer's disease; Huntington's disease; reproductive disorder; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             4..2532
FT                   /*tag= a
FT                   /product= "REPTR1 protein"
XX
PN   WO200198354-A2.
XX
PD   27-DEC-2001.
XX
PF   21-JUN-2001; 2001WO-US019942.
XX
PR   21-JUN-2000; 2000US-0214027P.
PR   25-AUG-2000; 2000US-0228045P.
PR   12-DEC-2000; 2000US-0255104P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Griffin JA,  Kallick DA,  Tribouley CM,  Yue H,  Nguyen DB,  Tang YT;
PI   Lal P,  Policky JL,  Azimzai Y,  Lu DAM,  Graul R,  Yao MG,  Burford N;
PI   Hafalia AJA,  Baughn MR,  Bandman O,  Patterson C,  Yang J,  Xu Y;
PI   Gandhi AR,  Warren BA,  Ding L,  Sanjanwala MS,  Duggan BM,  Lu Y;
XX
DR   WPI; 2002-090432/12.
DR   P-PSDB; AAU17818.
XX
PT   Twelve human receptors (referred to as REPTR-1 to REPTR-12), useful in
PT   the diagnosis, treatment and prevention of gastrointestinal (e.g.
PT   gastritis), autoimmune/inflammatory (e.g. osteoarthritis) and cell
PT   proliferative (e.g. cancer)disorders.
XX
PS   Claim 57; Page 142-143; 157pp; English.
XX
CC   This invention relates to twelve human receptors cDNA sequences referred
CC   to as REPTR-1 to REPTR-12), and the proteins encoded thereby. The
CC   proteins of the invention may have antiinflammatory, cytostatic,
CC   immunosuppressive, antiviral, anti-HIV, antiarthritic, muscular active
CC   general, anticonvulsant, nootropic, neuroprotective, antiallergic
CC   activities. The sequences of the invention may be used to produce REPTR
CC   agonists or antagonists, and the protein sequences may be used to raise
CC   anti-REPTR antibodies. These molecules and the REPTR polynucleotides and
CC   polypeptides of the invention are useful in the diagnosis,treatment and
CC   prevention of gastrointestinal (e.g. gastritis, oesophageal carcinoma,
CC   Crohn's disease, irritable bowel syndrome, ulcerative colitis), endocrine
CC   (e.g. hypothalamus disorder, Kallman's disease), autoimmune/ inflammatory
CC   (e.g. acquired immune deficiency syndrome (AIDS), rheumatoid arthritis,
CC   allergies, osteoarthritis, diabetes mellitus, multiple sclerosis,
CC   systemic lupus erythematosus), cell proliferative (e.g. cancer),
CC   developmental (e.g. Duchenne and Becker muscular dystrophy), neurological
CC   (e.g. epilepsy, Alzheimer's disease, Huntington's disease) and
CC   reproductive (e.g. infertility, endometriosis) disorders. Numerous other
CC   examples of each disorder are given in the specification. The present
CC   sequence represents the human REPTR1 cDNA sequence of the invention
XX
SQ   Sequence 3580 BP; 670 A; 1233 C; 1025 G; 652 T; 0 U; 0 Other;

  Query Match             83.4%;  Score 2402.6;  DB 6;  Length 3580;
  Best Local Similarity   93.4%;  Pred. No. 0;
  Matches 2615;  Conservative    0;  Mismatches    9;  Indels  177;  Gaps    3;

Qy          84 GCCATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGG 143
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 GCCATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGG 60

Qy         144 CTCCGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAAC 203
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 CTCCGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAAC 120

Qy         204 CCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAG 263
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 CCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAG 180

Qy         264 CCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGG 323
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 CCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGG 240

Qy         324 GAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAG 383
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||    |
Db         241 GAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTG 300

Qy         384 CCGACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTG 443
               || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 CCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTG 360

Qy         444 GAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAG 503
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 GAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAG 420

Qy         504 GCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAG 563
               ||||||||||||||||||    ||||||||||||||||||||||||||||||||||||||
Db         421 GCCTACATCCGCATAGCCTATTTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAG 480

Qy         564 GTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCC 623
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 GTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCC 540

Qy         624 GAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATAC 683
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 GAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATAC 600

Qy         684 ATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTAC 743
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 ATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTAC 660

Qy         744 ACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTC 803
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 ACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTC 720

Qy         804 TACGTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGG 863
               |||
Db         721 TAC--------------------------------------------------------- 723

Qy         864 CGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCT 923
```

```
Db        724 ------------------------------------------------------------ 723

Qy        924 TTCTGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGAC 983
                                                                    ||||||
Db        724 ------------------------------------------------------GTGGAC 729

Qy        984 GGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGG 1043
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        730 GGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGG 789

Qy       1044 AGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGAC 1103
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        790 AGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGAC 849

Qy       1104 CTGGACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGAC 1163
              |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db        850 CTGGACACCCGCAACTGTACCAGTGACCTCTGTGTACACACTGCTTCTGGCCCTGAGGAC 909

Qy       1164 GTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTC 1223
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        910 GTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTC 969

Qy       1224 CTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCC 1283
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        970 CTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCC 1029

Qy       1284 ATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCAT 1343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1030 ATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCAT 1089

Qy       1344 CTGCTCACCATCCAGCCGGACCTCAG---CACCACCACCACCTACCAGGGCAGTCTCTGT 1400
              ||||||||||||||||||||||||||   |||||||||||||||||||||||||||||||
Db       1090 CTGCTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGT 1149

Qy       1401 CCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCC 1460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1150 CCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCC 1209

Qy       1461 CTGGGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTC 1520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1210 CTGGGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTC 1269

Qy       1521 GTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATG 1580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1270 GTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATG 1329

Qy       1581 ACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGC 1640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db       1330 ACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGAATCAGC 1389

Qy       1641 CTCCTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTG 1700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1390 CTCCTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTG 1449

Qy       1701 CACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATC 1760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1450 CACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATC 1509

Qy       1761 GTTAGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCAC 1820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1510 GTTAGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCAC 1569

Qy       1821 TGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGC 1880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1570 TGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGC 1629

Qy       1881 AGCTGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGC 1940
              |||||||   ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1630 AGCTGGG---AGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGC 1686

Qy       1941 CAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTG 2000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1687 CAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTG 1746

Qy       2001 GGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTG 2060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1747 GGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTG 1806

Qy       2061 GCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCA 2120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1807 GCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCA 1866

Qy       2121 CTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGG 2180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1867 CTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGG 1926

Qy       2181 GTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGC 2240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1927 GTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGC 1986

Qy       2241 TCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGG 2300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1987 TCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGG 2046

Qy       2301 AATGGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACT 2360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2047 AATGGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACT 2106

Qy       2361 AGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGC 2420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2107 AGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGC 2166

Qy       2421 ATCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAA 2480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2167 ATCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAA 2226

Qy       2481 GCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAG 2540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2227 GCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAG 2286

Qy       2541 AAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCC 2600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2287 AAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCC 2346

Qy       2601 CAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCC 2660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2347 CAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCC 2406

Qy       2661 ATGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCT 2720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2407 ATGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCT 2466

Qy       2721 GCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAG 2780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2467 GCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAG 2526

Qy       2781 TGCTGAGGCCGGCCAGGCCCGACACCTACACTCTCACCAGCTTTGGCACCCACCAAGGAC 2840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2527 TGCTGAGGCCGGCCAGGCCCGACACCTACACTCTCACCAGCTTTGGCACCCACCAAGGAC 2586

Qy       2841 AGGCAGAAGCCGGACAGGGGCCCTTCCCCACACCGGGGAGA 2881
              |||||||||||||||||||||||||||||||||||||||||
Db       2587 AGGCAGAAGCCGGACAGGGGCCCTTCCCCACACCGGGGAGA 2627
```

```
RESULT 5
AAK52261
ID   AAK52261 standard; cDNA; 2907 BP.
XX
AC   AAK52261;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human polynucleotide SEQ ID NO 806.
XX
```

```
KW   Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW   vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW   tissue growth factor; immunomodulatory; cancer; leukaemia;
KW   nervous system disorder; arthritis; inflammation; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157190-A2.
XX
PD   09-AUG-2001.
XX
PF   05-FEB-2001; 2001WO-US004098.
XX
PR   03-FEB-2000; 2000US-00496914.
PR   27-APR-2000; 2000US-00560875.
PR   20-JUN-2000; 2000US-00598075.
PR   19-JUL-2000; 2000US-00620325.
PR   01-SEP-2000; 2000US-00654936.
PR   15-SEP-2000; 2000US-00663561.
PR   20-OCT-2000; 2000US-00693325.
PR   30-NOV-2000; 2000US-00728422.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Drmanac RT,  Asundi V,  Zhou P,  Xu C,  Cao Y;
PI   Ma Y,  Zhao QA,  Wang D,  Wang J,  Zhang J,  Ren F,  Chen R,  Wang ZW;
PI   Xue AJ,  Yang Y,  Wejhrman T,  Goodrich R;
XX
DR   WPI; 2001-476283/51.
DR   P-PSDB; AAM79128.
XX
PT   Nucleic acids encoding polypeptides with cytokine-like activities, useful
PT   in diagnosis and gene therapy.
XX
PS   Claim 1; Page 2691-2694; 6221pp; English.
XX
CC   The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC   encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
CC   cytokine, cell proliferation or cell differentiation or which may induce
CC   production of other cytokines in other cell populations. The
CC   polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC   peptide therapy. The polypeptides have various cytokine-like activities,
CC   e.g. stem cell growth factor activity, haematopoiesis regulating
CC   activity, tissue growth factor activity, immunomodulatory activity and
CC   activin/inhibin activity and may be useful in the diagnosis and/or
CC   treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC   inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
CC   (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
CC   sequence listing were missing at the time of publication
XX
SQ   Sequence 2907 BP; 552 A; 966 C; 881 G; 508 T; 0 U; 0 Other;

  Query Match            79.7%;  Score 2296.2;  DB 4;  Length 2907;
  Best Local Similarity  92.0%;  Pred. No. 0;
  Matches 2565;  Conservative    0;  Mismatches   38;  Indels  186;  Gaps    6;

Qy        133 TCGCCGCTTGGCTCCGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGC 192
Db        170 TGGCGGGCAGACAGCACGGATGGAGCGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGC 229

Qy        193 CTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCG 252
Db        230 CTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCG 289

Qy        253 TCAAGAACAAGCCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCA 312
Db        290 TCAAGAACAAGCCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCA 349

Qy        313 AGTGCAACGGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGA 372
Db        350 AGTGCAACGGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGA 409

Qy        373 GCAGTGGTGAGCCGACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGG 432
Db        410 GCAGTGGGCTGCCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGG 469

Qy        433 TGTTCGGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCA 492
Db        470 TGTTCGGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCA 529

Qy        493 AGAGTCAGAAGGCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGC 552
Db        530 AGAGTCAGAAGGCCTACATCCGCATAGCCTATTTGCGCAAGAACTTCGAGCAGGAGCCGC 589

Qy        553 TGGCCAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCA 612
Db        590 TGGCCAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCA 649

Qy        613 TCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACC 672
Db        650 TCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACC 709

Qy        673 CCAATGTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACA 732
Db        710 CCAATGTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACA 769

Qy        733 CGGCCAACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTG 792
Db        770 CGGCCAACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTG 829

Qy        793 CTGTCATCGTCTAC---------------------------------------------- 806
Db        830 CTGTCATCGTCTACGGTGGGCCCCGGGACTCCCTGGTCACAGGGAGAGGCACTGCGGTGC 889

Qy        807 -----------------------------------------GTGAACGGTGGGTGGTCGA 825
Db        890 CCCTGGGCAGTGACATGTGGCTGTCCTTCTCTGTCCGGCCAGTGAACGGTGGGTGGTCGA 949

Qy        826 CGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGGAGCC 885
Db        950 CGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGGAGCC 1009

Qy        886 GGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAATGTCC 945
Db       1010 GGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAATGTCC 1069

Qy        946 ATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGCAGCTGGAGCCCGTGGAGCA 1005
Db       1070 AGAA---AACAGCCTGCGCCACCCTGTGCCCAGTGGACGGCAGCTGGAGCCCGTGGAGCA 1126

Qy       1006 AGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCAG 1065
Db       1127 AGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCAG 1186

Qy       1066 CACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGTACCA 1125
Db       1187 CACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGTACCA 1246

Qy       1126 GTGACCTCTGTGTACACA------------------------------------------ 1143
Db       1247 GTGACCTCTGTGTACACAACTCCTACACCCCTGCCCCCACCAAGGCCATGCTGTCTCCCG 1306

Qy       1144 ---GTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCT 1200
Db       1307 CAGCTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCT 1366

Qy       1201 GCCTGGTCCTGCTGCTGCTTGTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGG 1260
Db       1367 GCCTGGTCCTGCTGCTGCTTGTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGG 1426

Qy       1261 ACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGC 1320
Db       1427 ACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGC 1486

Qy       1321 CCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCGGACCTCAG---CACCACCA 1377
```

```
Db       1487 CCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCA 1546

Qy       1378 CCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCA 1437
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1547 CCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCA 1606

Qy       1438 CCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCACACACTGCACCACAGCTCTC 1497
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1607 CCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCACACACTGCACCACAGCTCTC 1666

Qy       1498 CCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCC 1557
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1667 CCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCC 1726

Qy       1558 TGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGA 1617
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1727 TGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGA 1786

Qy       1618 TGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCCAGATGCCATACCCCGAGGGAAGA 1677
              |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db       1787 TGATCCCTAATACAGGAATCAGCCTCCTCATCCCCCCAGATGCCATACCCCGAGGGAAGA 1846

Qy       1678 TCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCT 1737
              |||||||||||||||||||||||||||||||||||||||||
Db       1847 TCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG------------------- 1887

Qy       1738 GTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGC 1797
                                        ||||||||||||||||||||||||||||||||||
Db       1888 --------------------------AGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGC 1921

Qy       1798 CAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCC 1857
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1922 CAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCC 1981

Qy       1858 TCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGG 1917
              ||||||||||||||||||||||||||||||   |||||||||||||||||||||||||||
Db       1982 TCAAAAAGCAGTCGTGCGAGGGCAGCTGGG---AGGATGTGCTGCACCTGGGCGAGGAGG 2038

Qy       1918 CGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGC 1977
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2039 CGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGC 2098

Qy       1978 AGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCA 2037
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2099 AGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCA 2158

Qy       2038 AGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACT 2097
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2159 AGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACT 2218

Qy       2098 GCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGG 2157
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2219 GCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGG 2278

Qy       2158 GACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCC 2217
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2279 GACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCC 2338

Qy       2218 TATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGG 2277
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2339 TATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGG 2398

Qy       2278 AGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTTGCACTGCACCTTCACCC 2337
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2399 AGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTTGCACTGCACCTTCACCC 2458

Qy       2338 TGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGG 2397
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2459 TGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGG 2518

Qy       2398 AGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACCAAGGACACAAGGTTTGCTG 2457
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2519 AGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACCAAGGACACAAGGTTTGCTG 2578

Qy       2458 AGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCA 2517
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2579 AGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCA 2638

Qy       2518 AGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGG 2577
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2639 AGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGG 2698

Qy       2578 GTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTG 2637
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2699 GTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTG 2758

Qy       2638 CCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCA 2697
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2759 CCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCA 2818

Qy       2698 ACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCC 2757
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2819 ACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCC 2878

Qy       2758 TCTTCACAGTGTCGGAGGCTGAGTGCTGA 2786
              |||||||||||||||||||||||||||||
Db       2879 TCTTCACAGTGTCGGAGGCTGAGTGCTGA 2907
```

```
RESULT 6
AAV52940
ID   AAV52940 standard; cDNA; 3014 BP.
XX
AC   AAV52940;
XX
DT   25-MAR-2003  (revised)
DT   21-DEC-1998  (first entry)
XX
DE   Rat UNC-5 homologue unc5h-1 cDNA.
XX
KW   UNC-5; UNC5H-1; rat; netrin receptor; cell migration; axon guidance;
KW   diagnosis; therapy; ds.
XX
OS   Rattus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..2697
FT                   /*tag=  a
XX
PN   WO9837085-A1.
XX
PD   27-AUG-1998.
XX
PF   19-FEB-1998;   98WO-US003143.
XX
PR   19-FEB-1997;   97US-00808982.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Tessier-Lavigne M,  Leonardo ED,  Hinck L,  Masu M,  Keinomasu K;
XX
DR   WPI; 1998-495364/42.
DR   P-PSDB; AAW78898.
XX
PT   Netrin-binding, vertebrate proteins - useful for diagnosis, therapy and
PT   the biopharmaceutical industry.
XX
PS   Claim 7; Page 15-17; 32pp; English.
XX
CC   This cDNA, termed unc5h-1, comprises a rat homologue of Caenorhabditis
CC   elegans unc-5. Rat unc5h-1 and unc5h-2 (see AAV52942) cDNAs were isolated
CC   from an E18 brain cDNA library. The predicted proteins (see AAW78898 and
CC   AAW78900) show similarity with UNC-5. They are predicted to be involved
CC   in cell migration and axon guidance, and are characterised as receptor
CC   proteins for netrins. Gene expression is observed in regions where
CC   differentiating neurons are undergoing axogenesis. Human unc5h-1 (see
```

```
CC   AAV52941) and unc5h-2 (see AAV52943) cDNAs are also claimed. Vertebrate
CC   UNC-5 proteins may be produced recombinantly from transfected host cells
CC   by utilising these vertebrate UNC-5 nucleic acids. The invention also
CC   provides unc-5 hybridisation probes and primers, vertebrate UNC-5-
CC   specific binding agents such as specific antibodies, and methods of
CC   making and using the subject compositions in diagnosis (e.g. genetic
CC   hybridisation screens for vertebrate unc-5 transcripts), therapy (e.g.
CC   gene therapy to modulate vertebrate unc-5 gene expression) and in the
CC   biopharmaceutical industry (e.g. as immunogens, reagents for modulating
CC   cell guidance, reagents for screening chemical libraries for lead
CC   pharmacological agents, etc.). (Updated on 25-MAR-2003 to correct PI
CC   field.)
XX
SQ   Sequence 3014 BP; 596 A; 977 C; 849 G; 592 T; 0 U; 0 Other;

  Query Match             77.6%;  Score 2235.8;  DB 2;  Length 3014;
  Best Local Similarity   88.5%;  Pred. No. 0;
  Matches 2474;  Conservative    0;  Mismatches  312;  Indels   11;  Gaps    4;

Qy         87 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 146
Db          1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGTGCTCCTGGGCATAGTCCTCGCCGCCTGGCTT 60

Qy        147 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 206
Db         61 CGTGGTTCGGGTGCCCAGCAGAGTGCCACGGTGGCCAATCCAGTGCCCGGTGCCAACCCC 120

Qy        207 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 266
Db        121 GACCTGCTGCCCCACTTCCTGGTAGAGCCTGAGGACGTGTACATTGTCAAGAACAAGCCG 180

Qy        267 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 326
Db        181 GTGTTGTTGGTGTGCAAGGCTGTGCCTGCCACCCAGATCTTCTTCAAGTGCAATGGGGAA 240

Qy        327 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCG 386
Db        241 TGGGTCCGCCAGGTCGATCACGTAATTGAACGCAGCACCGACAGCAGCAGCGGATTGCCA 300

Qy        387 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 446
Db        301 ACCATGGAGGTCCGTATCAACGTATCGAGGCAGCAGGTAGAGAAAGTGTTTGGGCTGGAG 360

Qy        447 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 506
Db        361 GAATACTGGTGCCAGTGTGTGGCATGGAGCTCCTCGGGTACCACCAAAAGTCAGAAGGCC 420

Qy        507 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 566
Db        421 TACATCCGGATTGCCTATTTGCGCAAGAACTTTGAGCAGGAGCCACTGGCCAAGGAAGTG 480

Qy        567 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 626
Db        481 TCACTGGAGCAAGGCATTGTACTACCTTGTCGCCCCCCAGAAGGAATCCCCCCAGCTGAG 540

Qy        627 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 686
Db        541 GTGGAGTGGCTTCGAAATGAGGACCTCGTGGACCCCTCCCTCGATCCCAATGTGTACATC 600

Qy        687 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 746
Db        601 ACGCGGGAGCACAGCCTAGTCGTGCGTCAGGCCCGCCTGGCCGACACGGCCAACTACACC 660

Qy        747 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 806
Db        661 TGTGTGGCCAAGAACATCGTAGCCCGTCGCCGAAGCACCTCTGCAGCGGTCATTGTTTAT 720

Qy        807 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 866
Db        721 GTGAACGGTGGGTGGTCGACGTGGACTGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGT 780

Qy        867 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 926
Db        781 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCACCTCTCAACGGGGGCGCCTTC 840

Qy        927 TGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGC 986
Db        841 TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACTCTG---TGCCCAGTGGATGGG 897

Qy        987 AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC 1046
Db        898 AGCTGGAGTTCGTGGAGTAAGTGGTCAGCCTGTGGGCTTGACTGCACCCACTGGCGGAGC 957

Qy       1047 CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG 1106
Db        958 CGCGAGTGCTCTGACCCAGCACCCCGCAATGGAGGTGAGGAGTGTCGGGGTGCTGACCTG 1017

Qy       1107 GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG 1166
Db       1018 GACACCCGCAACTGTACCAGTGACCTCTGCCTGCACACCGCTTCTTGCCCCGAGGACGTG 1077

Qy       1167 GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC 1226
Db       1078 GCTCTCTACATCGGCCTTGTCGCTGTGGCTGTGTGCCTCTTCTTGCTGTTGCTGGCCCTT 1137

Qy       1227 ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1286
Db       1138 GGACTCATTTACTGTCGCAAGAAGGAAGGGCTGGACTCCGATGTGGCCGACTCGTCCATC 1197

Qy       1287 CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1346
Db       1198 CTCACCTCGGGCTTCCAGCCTGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCACCTG 1257

Qy       1347 CTCACCATCCAGCCGGACCTCAGCACCACC---ACCACCTACCAGGGCAGTCTCTGTCCC 1403
Db       1258 CTCACCATCCAGCCAGACCTCAGCACCACCACTACCACCTACCAGGGCAGTCTATGTTCG 1317

Qy       1404 CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG 1463
Db       1318 AGGCAGGATGGACCCAGCCCCAAGTTCCAGCTCTCTAATGGTCACCTGCTCAGCCCACTG 1377

Qy       1464 GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC 1523
Db       1378 GGGAGTGGCCGCCATACGTTGCACCACAGCTCACCCACCTCTGAGGCTGAGGACTTCGTC 1437

Qy       1524 TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC 1583
Db       1438 TCCCGCCTCTCCACCCAAAACTACTTTCGTTCCCTGCCCCGCGGCACCAGCAACATGGCC 1497

Qy       1584 TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC 1643
Db       1498 TACGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACGGGGATCAGCCTC 1557

Qy       1644 CTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC 1703
Db       1558 CTCATACCCCCGGATGCCATCCCCCGAGGAAAGATCTACGAGATCTACCTCACACTGCAC 1617

Qy       1704 AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT 1763
Db       1618 AAGCCAGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCAGTCGTT 1677

Qy       1764 AGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT 1823
Db       1678 AGCTGTGGGCCCCCAGGAGTCCTGCTCACCCGGCCAGTCATCCTTGCAATGGACCACTGT 1737

Qy       1824 GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC 1883
Db       1738 GGAGAGCCCAGCCCTGACAGCTGGAGTCTGCGCCTCAAAAAGCAGTCCTGCGAGGGCAGT 1797

Qy       1884 TGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG 1943
Db       1798 TGGG---AGGATGTGCTGCACCTTGGTGAGGAGTCACCTTCCCACCTCTACTACTGCCAG 1854

Qy       1944 CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA 2003
Db       1855 CTGGAGGCCGGGGCCTGCTATGTCTTCACGGAGCAGCTGGGCCGCTTTGCCCTGGTAGGA 1914
```

```
Qy   2004 GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC 2063
Db   1915 GAGGCCCTCAGCGTGGCTGCCACCAAGCGCCTCAGGCTCCTTCTGTTTGCTCCCGTGGCC 1974

Qy   2064 TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC 2123
Db   1975 TGTACGTCCCTTGAGTACAACATCCGAGTGTACTGCCTACACGACACCCACGACGCTCTC 2034

Qy   2124 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTC 2183
Db   2035 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTAGGTGGACAGCTGATCCAGGAGCCTCGCGTC 2094

Qy   2184 CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCC 2243
Db   2095 CTGCACTTCAAAGACAGTTACCACAACCTACGTCTCTCCATCCACGACGTGCCCAGCTCC 2154

Qy   2244 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2303
Db   2155 CTGTGGAAGAGCAAGCTACTTGTCAGCTACCAGGAGATCCCTTTTTACCACATCTGGAAC 2214

Qy   2304 GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2363
Db   2215 GGCACCCAGCAGTATCTGCACTGCACCTTCACCCTGGAGCGCATCAACGCCAGCACCAGC 2274

Qy   2364 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2423
Db   2275 GACCTGGCCTGCAAGGTGTGGGTGTGGCAGGTGGAGGGAGATGGGCAGAGCTTCAACATC 2334

Qy   2424 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2483
Db   2335 AACTTCAACATCACTAAGGACACAAGGTTTGCTGAATTGTTGGCTCTGGAGAGTGAAGGG 2394

Qy   2484 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2543
Db   2395 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAAAAG 2454

Qy   2544 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2603
Db   2455 ATCATCGCCAGTCTGGACCCACCCTGCAGCCGGGGCGCCGACTGGAGAACTCTAGCCCAG 2514

Qy   2604 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG 2663
Db   2515 AAACTTCACCTGGACAGCCATCTTAGCTTCTTTGCCTCCAAGCCCAGCCCTACAGCCATG 2574

Qy   2664 ATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2723
Db   2575 ATCCTCAACCTATGGGAGGCACGGCACTTCCCCAACGGCAACCTCGGCCAGCTGGCAGCA 2634

Qy   2724 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2783
Db   2635 GCTGTGGCCGGACTGGGCCAACCAGATGCTGGCCTCTTCACGGTGTCGGAGGCCGAGTGT 2694

Qy   2784 TGAGGCCGGCCAGGCCCG--ACACCTACACTCTCACCAGCTTTGGCACCCACCAAGGACA 2841
Db   2695 TGAGACCAGCCAGGCCGGTAATGCCTACATTCTCACCAGCTTTGACACCTGCCAGGGACA 2754

Qy   2842 GGCAGAAGCCGGACAGGGGCCCTTCCCCACACCGGGG 2878
Db   2755 GGCAAAACCAGACAGGGGCCCTTCCCCCACACCCGGG 2791
```

RESULT 7
AAS16843
ID   AAS16843 standard; cDNA; 2697 BP.
XX
AC   AAS16843;
XX
DT   14-FEB-2002  (first entry)
XX
DE   Rat netrin receptor UNC5H1 (YSG7) cDNA.
XX
KW   YSG; YSG7; schizophrenia; chronic animal model; LCGU; netrin receptor;
KW   local cerebral glucose utilisation; phosphodiesterase 1-alpha; UNC5H1;
KW   calcium-independent alpha-latrotoxin receptor; CIRL; trkE; synapsin 1A;
KW   epithelial discoidin domain receptor 1; synapsin 1B; neuroleptic; ss;
KW   tumour necrosis factor alpha; TNF-alpha; rat.
XX
OS   Rattus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .2697
FT                   /*tag= a
FT                   /product= "Rat netrin receptor UNC5H1"
XX
PN   WO200175440-A2.
XX
PD   11-OCT-2001.
XX
PF   02-APR-2001; 2001WO-GB001486.
XX
PR   31-MAR-2000; 2000GB-00007880.
PR   26-MAY-2000; 2000GB-00012768.
XX
PA   (WELF-) WELFIDE CORP.
XX
PI   Cochran S, Paterson G, Ohashi Y, Morris B, Pratt J;
XX
DR   WPI; 2002-010813/01.
DR   P-PSDB; AAU10543.
XX
PT   Novel chronic animal model of schizophrenia, useful for identifying anti-
PT   psychotic drugs and genes that are associated with schizophrenia.
XX
PS   Claim 1; Fig 8a; 79pp; English.
XX
CC   The invention relates to YSG polynucleotide fragments for use in
CC   diagnosing and/or developing treatments for schizophrenia using chronic
CC   animal models. The polynucleotides and their encoded polypeptides are
CC   used for identification of compounds which modulate the expression of YSG
CC   molecules, leading to the manufacture of schizophrenia medicaments. The
CC   sequences can also be used for testing candidate compounds for any effect
CC   on the polypeptides. Anti-schizophrenic effects of a compound can be
CC   determined by measuring local cerebral glucose utilisation (LCGU) or
CC   comparing its expression level with that of a control group. The
CC   sequences are useful in the identification of genes associated with
CC   schizophrenic states and in the development of an antibody. The sequences
CC   of the invention include phosphodiesterase 1-alpha, calcium-independent
CC   alpha-latrotoxin receptors (CIRL)-1,2&3, epithelial discoidin domain
CC   receptor 1 (trkE), netrin receptor (UNC5H1), synapsins 1A and AB and
CC   tumour necrosis factor (TNF) alpha. This sequence represents rat netrin
CC   receptor UNC5H1 (YSG7) DNA
XX
SQ   Sequence 2697 BP; 541 A; 864 C; 766 G; 526 T; 0 U; 0 Other;

Query Match 76.0%; Score 2189.4; DB 6; Length 2697;
Best Local Similarity 88.9%; Pred. No. 0;
Matches 2403; Conservative 0; Mismatches 291; Indels 9; Gaps 3;

```
Qy     87 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 146
Db      1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGTGCTCCTGGGCATAGTCCTCGCCGCCTGGCTT 60

Qy    147 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 206
Db     61 CGTGGTTCGGGTGCCCAGCAGAGTGCCACGGTGGCCAATCCAGTGCCCGGTGCCAACCCC 120

Qy    207 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 266
Db    121 GACCTGCTGCCCCACTTCCTGGTAGAGCCTGAGGACGTGTACATTGTCAAGAACAAGCCG 180

Qy    267 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 326
Db    181 GTGTTGTTGGTGTGCAAGGCTGTGCCTGCCACCCAGATCTTCTTCAAGTGCAATGGGGAA 240

Qy    327 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCG 386
```

```
Db    241 TGGGTCCGCCAGGTCGATCACGTAATTGAACGCAGCACCGACAGCAGCAGCGGATTGCCA 300

Qy    387 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 446
Db    301 ACCATGGAGGTCCGTATCAACGTATCGAGGCAGCAGGTAGAGAAAGTGTTTGGGCTGGAG 360

Qy    447 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 506
Db    361 GAATACTGGTGCCAGTGTGTGGCATGGAGCTCCTCGGGTACCACCAAAAGTCAGAAGGCC 420

Qy    507 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 566
Db    421 TACATCCGGATTGCCTATTTGCGCAAGAACTTTGAGCAGGAGCCACTGGCCAAGGAAGTG 480

Qy    567 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 626
Db    481 TCACTGGAGCAAGGCATTGTACTACCTTGTCGCCCCCCAGAAGGAATCCCCCCAGCTGAG 540

Qy    627 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 686
Db    541 GTGGAGTGGCTTCGAAATGAGGACCTCGTGGACCCCTCCCTCGATCCCAATGTGTACATC 600

Qy    687 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 746
Db    601 ACGCGGGAGCACAGCCTAGTCGTGCGTCAGGCCCGCCTGGCCGACACGGCCAACTACACC 660

Qy    747 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 806
Db    661 TGTGTGGCCAAGAACATCGTAGCCCGTCGCCGAAGCACCTCTGCAGCGGTCATTGTTTAT 720

Qy    807 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 866
Db    721 GTGAACGGTGGGTGGTCGACGTGGACTGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGT 780

Qy    867 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 926
Db    781 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCACCTCTCAACGGGGGCGCCTTC 840

Qy    927 TGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGC 986
Db    841 TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACTCTG---TGCCCAGTGGATGGG 897

Qy    987 AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC 1046
Db    898 AGCTGGAGTTCGTGGAGTAAGTGGTCAGCCTGTGGGCTTGACTGCACCCACTGGCGGAGC 957

Qy   1047 CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG 1106
Db    958 CGCGAGTGCTCTGACCCAGCACCCCGCAATGGAGGTGAGGAGTGTCGGGGTGCTGACCTG 1017

Qy   1107 GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG 1166
Db   1018 GACACCCGCAACTGTACCAGTGACCTCTGCCTGCACACCGCTTCTTGCCCCGAGGACGTG 1077

Qy   1167 GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC 1226
Db   1078 GCTCTCTACATCGGCCTTGTCGCTGTGGCTGTGTGCCTCTTCTTGCTGTTGCTGGCCCTT 1137

Qy   1227 ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1286
Db   1138 GGACTCATTTACTGTCGCAAGAAGGAAGGGCTGGACTCCGATGTGGCCGACTCGTCCATC 1197

Qy   1287 CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1346
Db   1198 CTCACCTCGGGCTTCCAGCCTGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCACCTG 1257

Qy   1347 CTCACCATCCAGCCGGACCTCAGCACCACC---ACCACCTACCAGGGCAGTCTCTGTCCC 1403
Db   1258 CTCACCATCCAGCCAGACCTCAGCACCACCACTACCACCTACCAGGGCAGTCTATGTTCG 1317

Qy   1404 CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG 1463
Db   1318 AGGCAGGATGGACCCAGCCCCAAGTTCCAGCTCTCTAATGGTCACCTGCTCAGCCCACTG 1377

Qy   1464 GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC 1523
Db   1378 GGGAGTGGCCGCCATACGTTGCACCACAGCTCACCCACCTCTGAGGCTGAGGACTTCGTC 1437

Qy   1524 TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC 1583
Db   1438 TCCCGCCTCTCCACCCAAAACTACTTTCGTTCCCTGCCCCGCGGCACCAGCAACATGGCC 1497

Qy   1584 TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC 1643
Db   1498 TACGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACGGGGATCAGCCTC 1557

Qy   1644 CTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC 1703
Db   1558 CTCATACCCCCGGATGCCATCCCCCGAGGAAAGATCTACGAGATCTACCTCACACTGCAC 1617

Qy   1704 AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT 1763
Db   1618 AAGCCAGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCAGTCGTT 1677

Qy   1764 AGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT 1823
Db   1678 AGCTGTGGGCCCCCAGGAGTCCTGCTCACCCGGCCAGTCATCCTTGCAATGGACCACTGT 1737

Qy   1824 GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC 1883
Db   1738 GGAGAGCCCAGCCCTGACAGCTGGAGTCTGCGCCTCAAAAAGCAGTCCTGCGAGGGCAGT 1797

Qy   1884 TGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG 1943
Db   1798 TGGG---AGGATGTGCTGCACCTTGGTGAGGAGTCACCTTCCCACCTCTACTACTGCCAG 1854

Qy   1944 CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA 2003
Db   1855 CTGGAGGCCGGGGCCTGCTATGTCTTCACGGAGCAGCTGGGCCGCTTTGCCCTGGTAGGA 1914

Qy   2004 GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC 2063
Db   1915 GAGGCCCTCAGCGTGGCTGCCACCAAGCGCCTCAGGCTCCTTCTGTTTGCTCCCGTGGCC 1974

Qy   2064 TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC 2123
Db   1975 TGTACGTCCCTTGAGTACAACATCCGAGTGTACTGCCTACACGACACCCACGACGCTCTC 2034

Qy   2124 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTC 2183
Db   2035 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTAGGTGGACAGCTGATCCAGGAGCCTCGCGTC 2094

Qy   2184 CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCC 2243
Db   2095 CTGCACTTCAAAGACAGTTACCACAACCTACGTCTCTCCATCCACGACGTGCCCAGCTCC 2154

Qy   2244 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2303
Db   2155 CTGTGGAAGAGCAAGCTACTTGTCAGCTACCAGGAGATCCCTTTTTACCACATCTGGAAC 2214

Qy   2304 GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2363
Db   2215 GGCACCCAGCAGTATCTGCACTGCACCTTCACCCTGGAGCGCATCAACGCCAGCACCAGC 2274

Qy   2364 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2423
Db   2275 GACCTGGCCTGCAAGGTGTGGGTGTGGCAGGTGGAGGGAGATGGGCAGAGCTTCAACATC 2334

Qy   2424 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2483
Db   2335 AACTTCAACATCACTAAGGACACAAGGTTTGCTGAATTGTTGGCTCTGGAGAGTGAAGGG 2394

Qy   2484 GGGGTCCCAGCCCTGGTGGGCCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2543
Db   2395 GGGGTCCCAGCCCTGGTGGGCCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAAAAG 2454
```

```
Qy       2544 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2603
              || ||  |||| |||||||||||||| || ||||| ||||||||| | ||||| ||||||
Db       2455 ATCATCGCCAGTCTGGACCCACCCTGCAGCCGGGGCGCCGACTGGAGAACTCTAGCCCAG 2514

Qy       2604 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG 2663
              ||||| ||||||||||||||||| |||||||||||||||||||||||||| |||||||||
Db       2515 AAACTTCACCTGGACAGCCATCTTAGCTTCTTTGCCTCCAAGCCCAGCCCTACAGCCATG 2574

Qy       2664 ATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2723
              ||||||||||| |||||||| |||||||||||||||||||||||| |||||||||| |||
Db       2575 ATCCTCAACCTATGGGAGGCACGGCACTTCCCCAACGGCAACCTCGGCCAGCTGGCAGCA 2634

Qy       2724 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2783
              || ||||| ||||||||||| ||||| |||||||||||||| ||||||||||| ||||| 
Db       2635 GCTGTGGCCGGACTGGGCCAACCAGATGCTGGCCTCTTCACGGTGTCGGAGGCCGAGTGT 2694

Qy       2784 TGA 2786
              |||
Db       2695 TGA 2697


RESULT 8
AAV52941
ID   AAV52941 standard; cDNA; 1787 BP.
XX
AC   AAV52941;
XX
DT   25-MAR-2003  (revised)
DT   21-DEC-1998  (first entry)
XX
DE   Human UNC-5 homologue unc5h-1 cDNA.
XX
KW   UNC-5; UNC5H-1; human; netrin receptor; cell migration; axon guidance;
KW   diagnosis; therapy; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             3. .1670
FT                   /*tag=  a
FT                   /transl_except= (pos:21. .22, aa:Xaa)
FT                   /transl_except= (pos:200. .203, aa:Ile)
FT                   /transl_except= (pos:771. .772, aa:Ser)
FT                   /transl_except= (pos:785. .786, aa:Leu)
FT                   /transl_except= (pos:1078. .1079, aa:Xaa)
FT                   /transl_except= (pos:1098. .1099, aa:Xaa)
FT                   /transl_except= (pos:1106. .1107, aa:Xaa)
FT                   /transl_except= (pos:1621. .1622, aa:Gly)
FT                   /note= "these codons have apparent 1 nucleotide deletions
FT                   or insertions, which alter the reading frame"
XX
PN   WO9837085-A1.
XX
PD   27-AUG-1998.
XX
PF   19-FEB-1998;    98WO-US003143.
XX
PR   19-FEB-1997;    97US-00808982.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Tessier-Lavigne M,  Leonardo ED,  Hinck L,  Masu M,  Keinomasu K;
XX
DR   WPI; 1998-495364/42.
DR   P-PSDB; AAW78899.
XX
PT   Netrin-binding, vertebrate proteins - useful for diagnosis, therapy and
PT   the biopharmaceutical industry.
XX
PS   Claim 7; Page 17; 32pp; English.
XX
CC   This partial cDNA, termed unc5h-1, comprises a human homologue of
CC   Caenorhabditis elegans unc-5. unc5h-1 and unc5h-2 (see AAV52943) cDNAs
CC   were isolated from an embryonic brain cDNA library. The predicted
CC   proteins (see AAW78899 and AAW78901) show similarity with UNC-5. They are
CC   predicted to be involved in cell migration and axon guidance, and are
CC   characterised as receptor proteins for netrins. Gene expression is
CC   observed in regions where differentiating neurons are undergoing
CC   axogogenesis. Rat unc5h-1 (see AAV52940) and unc5h-2 (see AAV52942) cDNAs
CC   are also claimed. Vertebrate UNC-5 proteins may be produced recombinantly
CC   from transfected host cells by utilising these vertebrate UNC-5 nucleic
CC   acids. The invention also provides unc-5 hybridisation probes and
CC   primers, vertebrate UNC-5-specific binding agents such as specific
CC   antibodies, and methods of making and using the subject compositions in
CC   diagnosis (e.g. genetic hybridisation screens for vertebrate unc-5
CC   transcripts), therapy (e.g. gene therapy to modulate vertebrate unc-5
CC   gene expression) and in the biopharmaceutical industry (e.g. as
CC   immunogens, reagents for modulating cell guidance, reagents for screening
CC   chemical libraries for lead pharmacological agents, etc.). (Updated on 25
CC   -MAR-2003 to correct PI field.)
XX
SQ   Sequence 1787 BP; 349 A; 603 C; 491 G; 344 T; 0 U; 0 Other;

  Query Match             54.5%;  Score 1570.4;  DB 2;  Length 1787;
  Best Local Similarity   96.9%;  Pred. No. 5.6e-280;
  Matches 1721;  Conservative    0;  Mismatches   36;  Indels   19;  Gaps   11;

Qy       1114 GCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCT 1173
              |||||||||||||||||||||| ||||||| |||||||||||||||||||||||||||||
Db          1 GCAACTGTACCAGTGACCTCTG-GTACACACTGCTTCTGGCCCTGAGGACGTGGCCCTCT 59

Qy       1174 ATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCATCCTCG 1233
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         60 ATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCATCCTCG 119

Qy       1234 TTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCT 1293
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        120 TTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCT 179

Qy       1294 CAGGCTTCCAGCCCGTCAGCATC-AAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACC 1352
              ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        180 CAGGCTTCCAGCCCGTCAGCATCTAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACC 239

Qy       1353 ATCCAGCCGGACCTCAG---CACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAG 1409
              |||||||||||||||||   ||||||||||||||||||||||||||||||||||||||||
Db        240 ATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAG 299

Qy       1410 GATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGC 1469
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        300 GATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGC 359

Qy       1470 GGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGC 1529
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        360 GGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGC 419

Qy       1530 CTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGG 1589
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        420 CTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGG 479

Qy       1590 ACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATC 1649
              |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db        480 ACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGAATCAGCCTCCTCATC 539

Qy       1650 CCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCG 1709
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 CCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCG 599

Qy       1710 GAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGT 1769
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 GAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGT 659

Qy       1770 GGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAG 1829
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 GGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAG 719
```

```
Qy       1830 CCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAG 1889
Db        720 CCCAGCCCTGACAGCTGGAGCCTGGCCCTCAAAAAGCAGTCGTGCGAGGG-AGCTGGGAG 778

Qy       1890 CAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG 1949
Db        779 GATG----TCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG 834

Qy       1950 GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC 2009
Db        835 GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC 894

Qy       2010 CTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC 2069
Db        895 CTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC 954

Qy       2070 TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG 2129
Db        955 TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG 1014

Qy       2130 GTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC 2189
Db       1015 GTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC 1074

Qy       2190 TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG 2249
Db       1075 TT-AAGGACAGTTACCACAACCT--GCCCTATCATCCACGATGTGCCCAGCTCCCTGTGG 1131

Qy       2250 AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG 2309
Db       1132 AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG 1191

Qy       2310 CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG 2369
Db       1192 CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG 1251

Qy       2370 GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC 2429
Db       1252 GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC 1311

Qy       2430 AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC 2489
Db       1312 AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC 1371

Qy       2490 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT 2549
Db       1372 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT 1431

Qy       2550 TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC 2609
Db       1432 TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC 1491

Qy       2610 CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC 2669
Db       1492 CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC 1551

Qy       2670 AACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG 2729
Db       1552 AACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG 1611

Qy       2730 GCTGGACTGGGCCAGCCAGACGCTGGCCTC-TTCACAGTG-TCGGAGGCTGAGTGCTGAG 2787
Db       1612 GCTGGGACTGGCCAGCAGGACGGTGGCTTCTTTCACAGTGTTCGGAGGCTGAGTGCTGAG 1671

Qy       2788 GCCGGCCAGGCCCGACACCTACACTCTCACCAGCTTTGGCA--CCCACCAAGGACAGGCA 2845
Db       1672 GCCGGCCAGG--CGAACACTACAATTTTACCAGTTTTGGGAACCCACCAAGGGACAGGCA 1729

Qy       2846 GAAGCCGGACAGGGGCCCTTCCCCACACCGGGGAGA 2881
Db       1730 GAAGCCGGACAAGGGCTTTTCCCAAAACCGGGGAGA 1765
```

```
RESULT 9
AAC90958
ID   AAC90958 standard; cDNA; 9700 BP.
XX
AC   AAC90958;
XX
DT   19-MAR-2001  (first entry)
XX
DE   Plasmid pGC1037 nucleotide sequence SEQ ID NO:91.
XX
KW   Human; Caenorhabditis elegans; UNC-5; splice variant; nematode worm;
KW   protein-protein interaction; identification; ss.
XX
OS   Synthetic.
XX
PN   WO200073328-A2.
XX
PD   07-DEC-2000.
XX
PF   02-JUN-2000; 2000WO-EP005108.
XX
PR   01-JUN-1999;   99GB-00012755.
XX
PA   (DEVG-) DEVGEN NV.
XX
PI   Van Criekinge W,  Roelens I,  Bogaert T,  Verwaerde P;
XX
DR   WPI; 2001-016508/02.
XX
PT   Three variants of human unc-5C cDNAs (unc-5Cb, unc-5Cc and unc-5C8) and a
PT   human unc-5HS1 cDNA, useful in yeast two hybrid experiments for
PT   identifying unknown human cDNAs which encode proteins that interact with
PT   the human unc-5C protein.
XX
PS   Example 4; Page 228-230; 246pp; English.
XX
CC   The present invention describes 3 variants of human unc-5C cDNAs (unc-
CC   5Cb, unc-5Cc and unc-5C8) which correspond to alternatively spliced unc-
CC   5C transcripts, and a human unc-5HS1 cDNA which shares homology with the
CC   Rattus norvegicus unc-5HS1 cDNA. Also described are assays based on
CC   protein-protein-interactions between the unc-5 protein and a variety of
CC   different interacting proteins. The unc-5C variant cDNAs and unc-5HS1
CC   cDNA are useful in methods for identifying compounds which reduce or
CC   inhibit the lethal phenotype associated with the expression of the unc-5
CC   death domain in yeast. They are also useful in yeast two hybrid
CC   experiments for identifying unknown human cDNAs which encode proteins
CC   that interact with the human unc-5C protein. AAC90914 to AAC90971 and
CC   AAB50646 to AAB50693 represent sequences used in the exemplification of
CC   the present invention
XX
SQ   Sequence 9700 BP; 2604 A; 2193 C; 2236 G; 2666 T; 0 U; 1 Other;

  Query Match             44.8%;  Score 1291.2;  DB 4;  Length 9700;
  Best Local Similarity   99.5%;  Pred. No. 1.8e-228;
  Matches 1306;  Conservative    0;  Mismatches    3;  Indels    3;  Gaps    1;

Qy       1478 CACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCAC 1537
Db          7 CACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCAC 66

Qy       1538 CCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAA 1597
Db         67 CCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAA 126

Qy       1598 CTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCCAGA 1657
Db        127 CTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGAATCAGCCTCCTCATCCCCCCAGA 186

Qy       1658 TGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGT 1717
Db        187 TGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGT 246

Qy       1718 GAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCC 1777
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          247 GAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCC 306

Qy         1778 TGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCC 1837
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          307 TGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCC 366

Qy         1838 TGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGCAGGATGT 1897
                ||||||||||||||||||||||||||||||||||||||||||||||||||   |||||||
Db          367 TGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGG---AGGATGT 423

Qy         1898 GCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGC 1957
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          424 GCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGC 483

Qy         1958 CTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGT 2017
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          484 CTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGT 543

Qy         2018 GGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGA 2077
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          544 GGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGA 603

Qy         2078 GTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCA 2137
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          604 GTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCA 663

Qy         2138 GCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGA 2197
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          664 GCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGA 723

Qy         2198 CAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAA 2257
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          724 CAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAA 783

Qy         2258 GCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTA 2317
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          784 GCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTA 843

Qy         2318 CTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAA 2377
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          844 CTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAA 903

Qy         2378 GCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCAC 2437
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          904 GCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCAC 963

Qy         2438 CAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCT 2497
                ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db          964 CAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCAAGCCCT 1023

Qy         2498 GGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCT 2557
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1024 GGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCT 1083

Qy         2558 GGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGA 2617
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1084 GGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGA 1143

Qy         2618 CAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTG 2677
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1144 CAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTG 1203

Qy         2678 GGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACT 2737
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1204 GGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACT 1263

Qy         2738 GGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGAGGC 2789
                |||||||||||||||||||||||||||||||||||||||||||||||||| |
Db         1264 GGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGAGTC 1315
```

RESULT 10

```
AAH99530
ID   AAH99530 standard; cDNA; 1321 BP.
XX
AC   AAH99530;
XX
DT   16-OCT-2001  (first entry)
XX
DE   Human protein encoding cDNA sequence SEQ ID NO:365.
XX
KW   Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
KW   antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;
KW   antibacterial; endocrine; cardiant; central nervous system; virucide;
KW   anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;
KW   antiaggregant; haemostatic; vulnerary; antiulcer; osteopathic; eczema;
KW   dermatological; antiallergic; antiasthmatic; antidiabetic; cytostatic;
KW   neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
KW   immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
KW   antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
KW   cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
KW   genetic disease; haematopoietic disorder; platelet disorder; asthma;
KW   thrombocytopaenia; osteoporosis; severe combined immunodeficiency;
KW   allergic rhinitis; diabetes; multiple sclerosis; depression;
KW   Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
KW   neurological disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200153455-A2.
XX
PD   26-JUL-2001.
XX
PF   22-DEC-2000; 2000WO-US035017.
XX
PR   23-DEC-1999;   99US-00471275.
PR   21-JAN-2000; 2000US-00488725.
PR   25-APR-2000; 2000US-00552317.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Drmanac RT;
XX
DR   WPI; 2001-457603/49.
DR   P-PSDB; AAM25589.
XX
PT   Isolated human polynucleotides encoding polypeptides, useful for the
PT   treatment and diagnosis of e.g. cancer, ulcers and HIV infection.
XX
PS   Claim 1; Page 471; 1217pp; English.
XX
CC   AAH99166 to AAH99904 encode the human proteins given in AAM25225 to
CC   AAM25963. The proteins can have activities based on the tissues and cells
CC   they are expressed in, such as: antiinflammatory; antirheumatic;
CC   antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant;
CC   central nervous system; virucide; anti-HIV; fungicide; antimutagen;
CC   cardiovascular; antianaemic; antiaggregant; haemostatic; vulnerary;
CC   antiulcer; osteopathic; dermatological; antiallergic; antiasthmatic;
CC   antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;
CC   antiparkinsonian; and immunostimulant. The proteins and polynucleotides
CC   encoding them can be used in gene therapy, antisense therapy and vaccine
CC   production, The proteins and polynucleotides are useful for screening for
CC   agonists or antagonists of a protein and for the treatment and diagnosis
CC   of disorders associated with the activity of a protein e.g. inflammation,
CC   rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
CC   neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
CC   infections, autoimmunity, genetic diseases, haematopoietic disorders,
CC   anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,
CC   osteoporosis, severe combined immunodeficiency, eczema, allergic
CC   rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
CC   Alzheimer's disease, Parkinson's disease, neurodegenerative and
CC   neurological disorders
XX
SQ   Sequence 1321 BP; 258 A; 440 C; 371 G; 252 T; 0 U; 0 Other;
```

```
Query Match             41.4%;  Score 1193.6;  DB 4;  Length 1321;
Best Local Similarity   97.8%;  Pred. No. 1.4e-210;
Matches 1295;  Conservative    0;  Mismatches   19;  Indels   10;  Gaps    8;

Qy       1476 CACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCC 1535
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCC 60

Qy       1536 ACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTC 1595
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTC 120

Qy       1596 AACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCCA 1655
              |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db        121 AACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGAATCAGCCTCCTCATCCCCCCA 180

Qy       1656 GATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGAC 1715
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGAC 240

Qy       1716 GTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCC 1775
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCC 300

Qy       1776 CCT-GGCGTCCTGCTCACCCGGCCAGTCATCCT-GGCTATGGACCACTGT-GGGGAGCCC 1832
              ||| ||||||||||| ||||||||||||||||| || ||||||||||||| |||||||||
Db        301 CCTGGGCGTCCTGCTTACCCGGCCAGTCATCCTGGGGTATGGACCACTGTGGGGGAGCCC 360

Qy       1833 AGCCCTGACAGCT-GGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGCA 1891
              ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||   |
Db        361 AGCCCTGACAGCTGGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGG---A 417

Qy       1892 GGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGC 1951
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 GGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGC 477

Qy       1952 CAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCT 2011
              ||||||||||||||||||||||||||||||| |||||| |||||||||||||||||||||
Db        478 CAGTGCCTGCTACGTCTTCACCGAGCAGCTGAGCCGCTATGCCCTGGTGGGAGAGGCCCT 537

Qy       2012 CAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTC 2071
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        538 CAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTC 597

Qy       2072 CCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGT 2131
              ||||||||||||||| | ||||||||||||||||||||| |||||||||||||| |  ||
Db        598 CCTCGAGTACAACATACTGGTCTACTGCCTGCATGACACTCACGATGCACTCAACGTAGT 657

Qy       2132 GGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTT 2191
              ||||||||||||||||||||||  ||||||||||||||||||||||| || |||||||||
Db        658 GGTGCAGCTGGAGAAGCAGCTGCAGGGACAGCTGATCCAGGAGCCACTGGTACTGCACTT 717

Qy       2192 CAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAA 2251
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        718 CAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAA 777

Qy       2252 GAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCA 2311
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        778 GAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCA 837

Qy       2312 GCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGC 2371
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        838 GCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGC 897

Qy       2372 CTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAA 2431
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        898 CTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAA 957

Qy       2432 CATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCC 2491
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        958 CATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCC 1017

Qy       2492 AGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTC 2551
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1018 AGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTC 1077

Qy       2552 CAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCA 2611
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1078 CAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCA 1137

Qy       2612 CCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAA 2671
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1138 CCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAA 1197

Qy       2672 CCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGC 2731
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1198 CCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGC 1257

Qy       2732 T-GGACTGGGCCAGCCAGACGCTGGCCTC-TTCACAGTG-TCGGAGGCTGAGTGCTGAGG 2788
              | |||||||||||||  |||| |||| || ||||||||| ||||||||||||||||||||
Db       1258 TGGGACTGGGCCAGCAGGACGGTGGCTTCTTTCACAGTGTTCGGAGGCTGAGTGCTGAGG 1317

Qy       2789 CCGG 2792
              ||||
Db       1318 CCGG 1321
```

```
RESULT 11
AAS75738
ID   AAS75738 standard; cDNA; 3646 BP.
XX
AC   AAS75738;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #11542.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG11551.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 11542; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
```

```
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 3646 BP; 930 A; 917 C; 921 G; 877 T; 0 U; 1 Other;

  Query Match             31.8%;  Score 916;  DB 5;  Length 3646;
  Best Local Similarity   60.3%;  Pred. No. 2.1e-159;
  Matches 1657;  Conservative    1;  Mismatches 1016;  Indels   72;  Gaps    6;

Qy         98 GCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGG 157
Db        220 GCTCGTGCTACCTGCCCTGGCCCTGCTCAGCGCCAGCGGCACTGGCTCCGCCGCCCAAGA 279

Qy        158 TGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCC 217
Db        280 TGATGACTTTTTTCATGAACTCCCAGAAACTTTTCCTTCTGATCCACCTGAGCCTCTGCC 339

Qy        218 CCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGT 277
Db        340 ACATTTCCTTATTGAGCCTGAAGAAGCTTATATTGTGAAGAATAAGCCTGTGAACCTGTA 399

Qy        278 GTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCA 337
Db        400 CTGTAAAGCAAGCCCTGCCACCCAGATCTATTTCAAGTGTAATAGTGAATGGGTTCATCA 459

Qy        338 GGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCGACCATGGAGGT 397
Db        460 GAAGGACCACATAGTAGATGAAAGAGTAGATGAAACTTCCGGTCTCATTGTCCGGGAAGT 519

Qy        398 CCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTG 457
Db        520 GAGCATTGAGATTTCGCGCCAGCAAGTGGAAGAACTCTTTGGACCTGAAGATTACTGGTG 579

Qy        458 CCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCAT 517
Db        580 CCAGTGTGTGGCCTGGAGCTCCGCGGGTACCACAAAGAGCCGGAAGGCGTATGTGCGCAT 639

Qy        518 AGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCA 577
Db        640 TGCATATCTACGGAAGACATTTGAGCAGGAACCCCTAGGAAAGGAAGTGTCTTTGGAACA 699

Qy        578 GGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCT 637
Db        700 GGAAGTCTTACTCCAGTGTCGACCACCTGAAGGGATCCCAGTGGCTGAGGTGGAATGGTT 759

Qy        638 CCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCA 697
Db        760 GAAAAATGAAGACATAATTGATCCCGTTGAAGATCGGAATTTTTATATTACTATTGATCA 819

Qy        698 CAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAA 757
Db        820 CAACCTCATCATAAAGCAGGCCCGACTCTCTGATACTGCAAATTACACCTGTGTTGCCAA 879

Qy        758 GAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTACGTGAACGGTGG 817
Db        880 AAACATTGTTGCCAAGAGGAAAAGTACAACTGCCACTGTCATAGTCTATGTCAACGGTGG 939

Qy        818 GTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAA 877
Db        940 CTGGTCCACCTGGACGGAGTGGTCTGTGTGTAACAGCCGCTGTGGACGAGGGTATCAGAA 999

Qy        878 ACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCA 937
Db       1000 ACGTACAAGGACTTGTACCAACCCGGCACCACTCAATGGGGGTGCCTTCTGTGAAGGGCA 1059

Qy        938 GAATGTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGCAGCTGGAGCCC 997
Db       1060 GAGTGTGCAGAA---AATAGCCTGTACTACGTTATGCCCAGTGGATGGCAGGTGGACGCC 1116

Qy        998 GTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTC 1057
Db       1117 ATGGAGCAAGTGGTCTACTTGTGGAACTGAGTGCACCCACTGGCGCAGGAGGGAGTGCAC 1176

Qy       1058 TGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAA 1117
Db       1177 GGCGCCAGCCCCCAAGAATGGAGGCAAGGACTGCGACGGCCTCGTCTTGCAATCCAAGAA 1236

Qy       1118 CTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGT 1177
Db       1237 CTGCACTGATGGGCTTTGCATGCAGACTGCTCCTGATTCAGATGATGTTGCTCTCTATGT 1296

Qy       1178 GGG---CCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCATCCTCGT 1234
Db       1297 TGGGATTGTGATAGCAGTGATCGTTTGCCTGGCGATCTCTGTAGTTGTGGCCTTGTTTGT 1356

Qy       1235 TTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTC 1294
Db       1357 GTATCGGAAGAATCATCGTGACTTTGAGTCAGATATTATTGACTCTTCGGCACTCAATGG 1416

Qy       1295 AGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCAT 1354
Db       1417 GGGCTTTCARCCTGTGAACATCAAG---------GCAGCAAGACAAGATCTGCTGGCTGT 1467

Qy       1355 CCAGCCGGACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGG 1414
Db       1468 ACCCCCAGACCTCACGTCAGCTGCAGCCATGTACAGAGGACCTGTCTATGCCCTGCATGA 1527

Qy       1415 GCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCG 1474
Db       1528 CGTCTCAGACAAAATCCCAATGACCAACTCTCCAATTCTGGATCCACTGCCCAACCTGAA 1587

Qy       1475 CCACACACTGCACCACA-----------GCTCTCCCACCTCTGAGGCCGAGGAGTTCGT 1522
Db       1588 AATCAAAGTGTACAACACCTCAGGTGCTGTCTCCCCCCAAGATGACCTCTCTGAGTTTAC 1647

Qy       1523 CTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGA- 1581
Db       1648 GTCCAAGCTGTCCCCTCAGATGACCCAGTCGTTGTTGGAGAATGAAGCCCTCAGCCTGAA 1707

Qy       1582 -----------------------------------------CCTATGGGACCTTCAACTT 1600
Db       1708 GAACCAGAGTCTAGCAAGGCAGACTGATCCATCCTGTACCGCATTTGGCAGCTTCAACTC 1767

Qy       1601 CCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCCAGATGC 1660
Db       1768 GCTGGGAGGTCACCTTATTGTTCCCAATTCAGGAGTCAGCTTGCTGATTCCCGCTGGGGC 1827

Qy       1661 CATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTGAG 1720
Db       1828 CATTCCCCAAGGGAGAGTCTACGAAATGTATGTGACTGTACACAGGAAAGAAACTATGAG 1887

Qy       1721 GTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCTGG 1780
Db       1888 GCCACCCATGGATGACTCTCAGACACTTTTGACCCCTGTGGTGAGCTGTGGGCCCCCAGG 1947

Qy       1781 CGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCTGA 1840
Db       1948 AGCTCTGCTCACCCGCCCCGTCGTCCTCACTATGCATCACTGCGCAGACCCCAATACCGA 2007

Qy       1841 CAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGCAGGATGTGCT 1900
Db       2008 GGACTGGAAAATACTGCTCAAGAACCAGGCAGCACAGGGACAGTGGG---AGGATGTGGT 2064

Qy       1901 GCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTG 1960
Db       2065 GGTGGTCGGGGAGGAAAACTTCACCACCCCCTGCTACATTAAGCTGGATGCAGAGGCCTG 2124

Qy       1961 CTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGC 2020
```

```
Db       2125 CCACATCCTCACAGAGAACCTCAGCACCTACGCCCTGGTAGGACATTCCACCACCAAAGC 2184

Qy       2021 TGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAGTA 2080
Db       2185 GGCTGCAAAGCGCCTCAAGCTGGCCATCTTTGGGCCCCTGTGCTGCTCCTCGCTGGAGTA 2244

Qy       2081 CAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCT 2140
Db       2245 CAGCATCCGAGTCTACTGTCTGGATGACACCCAGGATGCCCTGAAGGAAATTTTACATCT 2304

Qy       2141 GGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAG 2200
Db       2305 TGAGAGACAGACGGGAGGACAGCTCCTAGAAGAACCTAAGGCTCTTCATTTTAAAGGCAG 2364

Qy       2201 TTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAGCT 2260
Db       2365 CACCCACAACCTGCGCCTGTCAATTCACGATATCGCCCATTCCCTCTGGAAGAGCAAATT 2424

Qy       2261 CCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTT 2320
Db       2425 GCTGGCTAAATATCAGGAAATTCCATTTTACCATGTTTGGAGTGGATCTCAAAGAAACCT 2484

Qy       2321 GCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAGCT 2380
Db       2485 GCACTGCACCTTCACTCTGGAAAGATTTAGCCTGAACACAGTGGAGCTGGTTTGCAAACT 2544

Qy       2381 GTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACCAA 2440
Db       2545 CTGTGTGCGGCAGGTGGAAGGAGAAGGGCAGATCTTCCAGCTCAACTGCACCGTGTCAGA 2604

Qy       2441 GGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTGGT 2500
Db       2605 GGAACCTACTGGCATCGATTTGCCGCTGCTGGATCCTGCGAACACCATCACCACGGTCAC 2664

Qy       2501 GGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTGGA 2560
Db       2665 GGGGCCCAGTGCTTTCAGCATCCCTCTCCCTATCCGGCAGAAGCTCTGTAGCAGCCTGGA 2724

Qy       2561 CCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGACAG 2620
Db       2725 TGCCCCCCAGACGAGAGGCCATGACTGGAGGATGCTGGCCCATAAGCTGAACCTGGACAG 2784

Qy       2621 CCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGGGA 2680
Db       2785 GTACTTGAATTACTTTGCCACCAAATCCAGCCCAACTGGCGTAATCCTGGATCTTTGGGA 2844

Qy       2681 GGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGG 2740
Db       2845 AGCACAGAACTTCCCAGATGGAAACCTGAGCATGCTGGCAGCTGTCTTGGAAGAAATGGG 2904

Qy       2741 CCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2786
Db       2905 AAGACATGAAACGGTGGTGTCCTTAGCAGCAGAAGGGCAGTATTAA 2950
```

RESULT 12
ABT06279

```
ID   ABT06279 standard; cDNA; 2860 BP.
XX
AC   ABT06279;
XX
DT   24-OCT-2002  (first entry)
XX
DE   Human NOV1a coding sequence.
XX
KW   Human; NOVX; autoimmune disease; cancer; infection; inflammatory disease;
KW   storage disorder; muscle disorder; neurodegenerative disorder; nootropic;
KW   developmental defect; neuroprotective; antiparkinsonian; hypotensive;
KW   hypertensive; haemostatic; cardiant; antianginal; dermatological;
KW   immunosuppressive; antiinflammatory; virucide; antibacterial; anti-HIV;
KW   antiparasitic; antiallergic; antiasthmatic; antirheumatic; antiarthritic;
KW   vulnerary; anorectic; antidiabetic; immunomodulator; antipsoriatic;
KW   nephrotropic; kerolytic; antiulcer; cerebroprotective; anticonvulsant;
KW   antiinfertility; antimanic; antidepressant; metabolic; cytostatic;
KW   tranquilizer; analgesic; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200257450-A2.
XX
PD   25-JUL-2002.
XX
PF   29-NOV-2001; 2001WO-US048922.
XX
PR   29-NOV-2000; 2000US-0253834P.
PR   30-NOV-2000; 2000US-0250926P.
PR   25-JAN-2001; 2001US-0264180P.
PR   20-AUG-2001; 2001US-0313656P.
PR   05-OCT-2001; 2001US-0327456P.
PR   28-NOV-2001; 2001US-00327456.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Edinger S,   Macdougall JR,   Millet I,   Ellerman K,   Stone DJ;
PI   Gerlach V,   Grosse WM,   Alsobrook JP,   Lepley DM,   Rieger D,   Burgess CE;
PI   Casman SJ,   Spytek KA,   Boldog FL,   Li L,   Padigaru M,   Mishra V;
PI   Patturajan M,   Shenoy S,   Rastelli L,   Tchernev VT,   Vernet CAM;
PI   Zerhusen BD,   Malyankar UM,   Guo X,   Miller CE,   Gangolli EA;
XX
DR   WPI; 2002-590741/63.
DR   P-PSDB; AAO18734.
XX
PT   Novel isolated polypeptide, designated NOVX, useful for treating or
PT   preventing in NOVX-associated disorders e.g. cardiomyopathy,
PT   atherosclerosis, diabetes, cancer, allergy, asthma, Crohn's disease.
XX
PS   Claim 9; Page 11; 353pp; English.
XX
CC   The present invention provides the protein and coding sequences of
CC   several novel human proteins, designated NOVX. These can be used in the
CC   treatment of, amongst others, cancers, autoimmune diseases, infections,
CC   inflammatory diseases, storage disorders, muscle disorders,
CC   neurodegenerative diseases and developmental defects. The present
CC   sequence is a coding sequence of the invention
XX
SQ   Sequence 2860 BP; 555 A; 950 C; 860 G; 495 T; 0 U; 0 Other;

  Query Match             30.9%;  Score 889;  DB 6;  Length 2860;
  Best Local Similarity   60.7%;  Pred. No. 1.9e-154;
  Matches 1637;  Conservative    0;  Mismatches  960;  Indels  102;  Gaps    7;

Qy        184 ACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATG 243
Db        168 ACTCCTTCCCGTCAGCGCCAGCAGAGCCGCTGCCCTACTTCCTGCAGGAGCCACAGGACG 227

Qy        244 TGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGA 303
Db        228 CCTACATTGTGAAGAACAAGCCTGTGGAGCTCCGCTGCCGCGCCTTCCCCGCCACACAGA 287

Qy        304 TCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCA 363
Db        288 TCTACTTCAAGTGCAACGGCGAGTGGGTCAGCCAGAACGACCACGTCACACAGGAAGGCC 347

Qy        364 CAGACGGGAGCAGTGGTGAGCCGACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGG 423
Db        348 TGGATGAGGCCACCGGTCTGCGGGTGCGCGAGGTGCAGATCGAGGTGTCGCGGCAGCAGG 407

Qy        424 TCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGG 483
Db        408 TGGAGGAGCTCTTTGGGCTGGAGGATTACTGGTGCCAGTGCGTGGCCTGGAGCTCCGCGG 467

Qy        484 GCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGC 543
Db        468 GCACCACCAAGAGTCGCCGAGCCTACGTCCGCATCGCCTACCTGCGCAAGAACTTCGATC 527
```

```
Qy    544 AGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCAC 603
Db    528 AGGAGCCTCTGGGCAAGGAGGTGCCCCTGGACCATGAGGTTCTCCTGCAGTGCCGCCCGC 587

Qy    604 CGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGT 663
Db    588 CGGAGGGGGTGCCTGTGGCCGAGGTGGAATGGCTCAAGAATGAGGATGTCATCGACCCCA 647

Qy    664 CCCTGGACCCCAATGTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCC 723
Db    648 CCCAGGACACCAACTTCCTGCTCACCATCGACCACAACCTCATCATCCGCCAGGCCCGCC 707

Qy    724 TTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCG 783
Db    708 TGTCGGACACTGCCAACTATACCTGCGTGGCCAAGAACATCGTGGCCAAACGCCGGAGCA 767

Qy    784 CCTCCGCTGCTGTCATCGTCTACGTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCG 843
Db    768 CCACTGCCACCGTCATCGTCTACGTGAATGGCGGCTGGTCCAGCTGGGCAGAGTGGTCAC 827

Qy    844 TCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGG 903
Db    828 CCTGCTCCAACCGCTGTGGCCGAGGCTGGCAGAAGCGCACCCGGACCTGCACCAACCCCG 887

Qy    904 CGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCCT 963
Db    888 CTCCACTCAACGGAGGGGCCTTCTGCGAGGGCCAGGCATTCCAGAA---GACCGCCTGCA 944

Qy    964 CTCTGCTTGTCTCTGTGGACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGC 1023
Db    945 CCACCATCTGCCCAGTCGATGGGGCGTGGACGGAGTGGAGCAAGTGGTCAGCCTGCAGCA 1004

Qy   1024 TGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGG 1083
Db   1005 CTGAGTGTGCCCACTGGCGTAGCCGCGAGTGCATGGCGCCCCCACCCCAGAACGGAGGCC 1064

Qy   1084 AGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGTACCAGTGACCTCTGTGTACACA 1143
Db   1065 GTGACTGCAGCGGGACGCTGCTCGACTCTAAGAACTGCACAGATGGGCTGTGCATGCAAC 1124

Qy   1144 GTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCC 1203
Db   1125 TGGAGGCCTCAGGGGATGCGGCGCTGTATGCGGGGCTCGTGGTGGCCATCTTCGTGGTCG 1184

Qy   1204 TGGTCCTGCTGCTGCTTGTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACT 1263
Db   1185 TGGCAATCCTCATGGCGGTGGGGGTGGTGGTGTACCGCCGCAACTGCCGTGACTTCGACA 1244

Qy   1264 CAGATGTGGCTGACTCGTCCATT---CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGC 1320
Db   1245 CAGACATCACTGACTCATCTGCTGCCCTGACTGGTGGTTTCCACCCCGTCAACTTTAAGA 1304

Qy   1321 CCAGCAAAGCAGACAACCCCCATCTGCT------CACCATCCAGCCGGACCTCAGCACCA 1374
Db   1305 CGGCAAGGCCCAGTAACCCGCAGCTCCTACACCCCTCTGTGCCTCCTGACCTGACAGCCA 1364

Qy   1375 CCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGC 1434
Db   1365 GCGCCGGCATCTACCGCGGACCCGTGTATGCCCTGCAGGACTCCACCGACAAAATCCCCA 1424

Qy   1435 TCACCAATGGGCACCTGCTCAGCCCC---------------------------------- 1460
Db   1425 TGACCAACTCTCCTCTGCTGGACCCCTTACCCAGCCTTAAGGTCAAGGTCTACAGCTCCA 1484

Qy   1461 --------------------------------------------CTGGGTGGCGGCCGCC 1476
Db   1485 GCACCACGGGCTCTGGGCCAGGCCTGGCAGATGGGGCTGACCTGCTGGGGGTCTTGCCGC 1544

Qy   1477 ACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCA 1536
Db   1545 CTGGCACATACCCTAGCGATTTCGCCCGGGACACCCACTTCCTGCACCTGCGCAGCGCCA 1604

Qy   1537 CCCAGAACTAC------TTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGA 1590
Db   1605 GCCTCGGTTCCCAGCAGCTCTTGGGCCTGCCCCGAGACCCAGGGAGCAGCGTCAGCGGCA 1664

Qy   1591 CCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCC 1650
Db   1665 CCTTTGGCTGCCTGGGTGGGAGGCTCAGCATCCCCGGCACAGGGGTCAGCTTGCTGGTGC 1724

Qy   1651 CCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGG 1710
Db   1725 CCAATGGAGCCATTCCCCAGGGCAAGTTCTACGAGATGTATCTACTCATCAACAAGGCAG 1784

Qy   1711 AAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTG 1770
Db   1785 AAAGTACCCTGCCGCTTTCAGAAGGGACCCAGACAGTATTGAGCCCCTCGGTGACCTGTG 1844

Qy   1771 GACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGC 1830
Db   1845 GACCCACAGGCCTCCTGCTGTGCCGCCCCGTCATCCTCACCATGCCCCACTGTGCCGAAG 1904

Qy   1831 CCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGC 1890
Db   1905 TCAGTGCCCGTGACTGGATCTTTCAGCTCAAGACCCAGGCCCACCAGGGCCACTGGGAG- 1963

Qy   1891 AGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGG 1950
Db   1964 --GAGGTGGTGACCCTGGATGAGGAGACCCTGAACACACCCTGCTACTGCCAGCTGGAGC 2021

Qy   1951 CCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCC 2010
Db   2022 CCAGGGCCTGTCACATCCTGCTGGACCAGCTGGGCACCTACGTGTTCACGGGCGAGTCCT 2081

Qy   2011 TCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCT 2070
Db   2082 ATTCCCGCTCAGCAGTCAAGCGGCTCCAGCTGGCCGTCTTCGCCCCCGCCCTCTGCACCT 2141

Qy   2071 CCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGG 2130
Db   2142 CCCTGGAGTACAGCCTCCGGGTCTACTGCCTGGAGGACACGCCTGTAGCACTGAAGGAGG 2201

Qy   2131 TGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACT 2190
Db   2202 TGCTGGAGCTGGAGCGGACTCTGGGCGGATACTTGGTGGAGGAGCCGAAACCGCTAATGT 2261

Qy   2191 TCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGA 2250
Db   2262 TCAAGGACAGTTACCACAACCTGCGCCTCTCCCTCCATGACCTCCCCCATGCCCATTGGA 2321

Qy   2251 AGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGC 2310
Db   2322 GGAGCAAGCTGCTGGCCAAATACCAGGAGATCCCCTTCTATCACATTTGGAGTGGCAGCC 2381

Qy   2311 AGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGG 2370
Db   2382 AGAAGGCCCTCCACTGCACTTTCACCCTGGAGAGGCACAGCTTGGCCTCCACAGAGCTCA 2441

Qy   2371 CCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCA 2430
Db   2442 CCTGCAAGATCTGCGTGCGGCAAGTGGAAGGGGAGGGCCAGATATTCCAGCTGCATACCA 2501

Qy   2431 ACATCACCAAG---GACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGG 2487
Db   2502 CTCTGGCAGAGACACCTGCTGGCTCCCTGGACACTCTCTGCTCTGCCCCTGGCAGCACTG 2561

Qy   2488 TCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAA 2547
Db   2562 TCACCACCCAGCTGGGACCTTATGCCTTCAAGATCCCACTGTCCATCCGCCAGAAGATAT 2621

Qy   2548 TTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAAC 2607
Db   2622 GCAACAGCCTAGATGCCCCCAACTCACGGGGCAATGACTGGCGGATGTTAGCACAGAAGC 2681

Qy   2608 TCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCC 2667
```

```
Db     2682 TCTCTATGGACCGGTACCTGAATTACTTTGCCACCAAAGCGAGCCCCACGGGTGTGATCC 2741

Qy     2668 TCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAG 2727

Db     2742 TGGACCTCTGGGAAGCTCTGCAGCAGGACGATGGGGACCTCAACAGCCTGGCGAGTGCCT 2801

Qy     2728 TGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2786

Db     2802 TGGAGGAGATGGGCAAGAGTGAGATGCTGGTGGCTGTGGCCACCGACGGGGACTGCTGA 2860
```

```
RESULT 13
ABT06280
ID   ABT06280 standard; cDNA; 2860 BP.
XX
AC   ABT06280;
XX
DT   24-OCT-2002  (first entry)
XX
DE   Human NOV1b coding sequence.
XX
KW   Human; NOVX; autoimmune disease; cancer; infection; inflammatory disease;
KW   storage disorder; muscle disorder; neurodegenerative disorder; nootropic;
KW   developmental defect; neuroprotective; antiparkinsonian; hypotensive;
KW   hypertensive; haemostatic; cardiant; antianginal; dermatological;
KW   immunosuppressive; antiinflammatory; virucide; antibacterial; anti-HIV;
KW   antiparasitic; antiallergic; antiasthmatic; antirheumatic; antiarthritic;
KW   vulnerary; anorectic; antidiabetic; immunomodulator; antipsoriatic;
KW   nephrotropic; kerolytic; antiulcer; cerebroprotective; anticonvulsant;
KW   antiinfertility; antimanic; antidepressant; metabolic; cytostatic;
KW   tranquilizer; analgesic; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200257450-A2.
XX
PD   25-JUL-2002.
XX
PF   29-NOV-2001; 2001WO-US048922.
XX
PR   29-NOV-2000; 2000US-0253834P.
PR   30-NOV-2000; 2000US-0250926P.
PR   25-JAN-2001; 2001US-0264180P.
PR   20-AUG-2001; 2001US-0313656P.
PR   05-OCT-2001; 2001US-0327456P.
PR   28-NOV-2001; 2001US-00327456.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Edinger S,  Macdougall JR,  Millet I,  Ellerman K,  Stone DJ;
PI   Gerlach V,  Grosse WM,  Alsobrook JP,  Lepley DM,  Rieger D,  Burgess CE;
PI   Casman SJ,  Spytek KA,  Boldog FL,  Li L,  Padigaru M,  Mishra V;
PI   Patturajan M,  Shenoy S,  Rastelli L,  Tchernev VT,  Vernet CAM;
PI   Zerhusen BD,  Malyankar UM,  Guo X,  Miller CE,  Gangolli EA;
XX
DR   WPI; 2002-590741/63.
DR   P-PSDB; AAO18735.
XX
PT   Novel isolated polypeptide, designated NOVX, useful for treating or
PT   preventing in NOVX-associated disorders e.g. cardiomyopathy,
PT   atherosclerosis, diabetes, cancer, allergy, asthma, Crohn's disease.
XX
PS   Claim 9; Page 14; 353pp; English.
XX
CC   The present invention provides the protein and coding sequences of
CC   several novel human proteins, designated NOVX. These can be used in the
CC   treatment of, amongst others, cancers, autoimmune diseases, infections,
CC   inflammatory diseases, storage disorders, muscle disorders,
CC   neurodegenerative diseases and developmental defects. The present
CC   sequence is a coding sequence of the invention
XX
SQ   Sequence 2860 BP; 557 A; 950 C; 858 G; 495 T; 0 U; 0 Other;
```

```
  Query Match             30.8%;  Score 887.4;  DB 6;  Length 2860;
  Best Local Similarity   60.6%;  Pred. No. 3.8e-154;
  Matches 1636;  Conservative    0;  Mismatches  961;  Indels  102;  Gaps    7;

Qy      184 ACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATG 243

Db      168 ACTCCTTCCCGTCAGCGCCAGCAGAGCCGCTGCCCTACTTCCTGCAGGAGCCACAGGACG 227

Qy      244 TGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGA 303

Db      228 CCTACATTGTGAAGAACAAGCCTGTGGAGCTTCGCTGCCGCGCCTTCCCCGCCACACAGA 287

Qy      304 TCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCA 363

Db      288 TCTACTTCAAGTGCAACGGCGAGTGGGTCAGCCAGAACGACCACGTCACACAGGAAGGCC 347

Qy      364 CAGACGGGAGCAGTGGTGAGCCGACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGG 423

Db      348 TGGATGAGGCCACCGGCCTGCGGGTGCGCGAGGTGCAGATCGAGGTGTCGCGGCAGCAGG 407

Qy      424 TCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGG 483

Db      408 TGGAGGAGCTCTTTGGGCTGGAGGATTACTGGTGCCAGTGCGTGGCCTGGAGCTCCGCAG 467

Qy      484 GCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGC 543

Db      468 GCACCACCAAGAGTCGCCGAGCCTACGTCCGCATCGCCTACCTGCGCAAGAACTTCGATC 527

Qy      544 AGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCAC 603

Db      528 AGGAGCCTCTGGGCAAGGAGGTGCCCCTGGACCATGAGGTTCTCCTGCAGTGCCGCCCGC 587

Qy      604 CGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGT 663

Db      588 CGGAGGGGGTGCCTGTGGCCGAGGTGGAATGGCTCAAGAATGAGGATGTCATCGACCCCA 647

Qy      664 CCCTGGACCCCAATGTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCC 723

Db      648 CCCAGGACACCAACTTCCTGCTCACCATCGACCACAACCTCATCATCCGCCAGGCCCGCC 707

Qy      724 TTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCG 783

Db      708 TGTCGGACACTGCCAACTATACCTGCGTGGCCAAGAACATCGTGGCCAAACGCCGGAGCA 767

Qy      784 CCTCCGCTGCTGTCATCGTCTACGTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCG 843

Db      768 CCACTGCCACCGTCATCGTCTACGTGAATGGCGGCTGGTCCAGCTGGGCAGAGTGGTCAC 827

Qy      844 TCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGG 903

Db      828 CCTGCTCCAACCGCTGTGGCCGAGGCTGGCAGAAGCGCACCCGGACCTGCACCAACCCCG 887

Qy      904 CGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCCT 963

Db      888 CTCCACTCAACGGAGGGGCCTTCTGCGAGGGCCAGGCATTCCAGAA---GACCGCCTGCA 944

Qy      964 CTCTGCTTGTCTCTGTGGACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGC 1023

Db      945 CCACCATCTGCCCAGTCGATGGGGCGTGGACGGAGTGGAGCAAGTGGTCAGCCTGCAGCA 1004

Qy     1024 TGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGG 1083

Db     1005 CTGAGTGTGCCCACTGGCGTAGCCGCGAGTGCATGGCGCCCCCACCCCAGAACGGAGGCC 1064

Qy     1084 AGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGTACCAGTGACCTCTGTGTACACA 1143

Db     1065 GTGACTGCAGCGGGACGCTGCTCGACTCTAAGAACTGCACAGATGGGCTGTGCATGCAAC 1124

Qy     1144 GTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCC 1203

Db     1125 TGGAGGCCTCAGGGGATGCGGCGCTGTATGCGGGGCTCGTGGTGGCCATCTTCGTGGTCG 1184

Qy     1204 TGGTCCTGCTGCTGCTTGTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACT 1263
```

```
Db   1185 TGGCAATCCTCATGGCGGTGGGGGTGGTGGTGTACCGCCGCAACTGCCGTGACTTCGACA 1244

Qy   1264 CAGATGTGGCTGACTCGTCCATT---CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGC 1320
Db   1245 CAGACATCACTGACTCATCTGCTGCCCTGACTGGTGGTTTCCACCCCGTCAACTTTAAGA 1304

Qy   1321 CCAGCAAAGCAGACAACCCCCATCTGCT------CACCATCCAGCCGGACCTCAGCACCA 1374
Db   1305 CGGCAAGGCCCAGTAACCCGCAGCTCCTACACCCCTCTGTGCCTCCTGACCTGACAGCCA 1364

Qy   1375 CCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGC 1434
Db   1365 GCGCCGGCATCTACCGCGGACCCGTGTATGCCCTGCAGGACTCCACCGACAAAATCCCCA 1424

Qy   1435 TCACCAATGGGCACCTGCTCAGCCCC-------------------------------- 1460
Db   1425 TGACCAACTCTCCTCTGCTGGACCCCTTACCCAGCCTTAAGGTCAAGGTCTACAGCTCCA 1484

Qy   1461 --------------------------------------------CTGGGTGGCGGCCGCC 1476
Db   1485 GCACCACGGGCTCTGGGCCAGGCCTGGCAGATGGGGCTGACCTGCTGGGGGTCTTGCCGC 1544

Qy   1477 ACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCA 1536
Db   1545 CTGGCACATACCCTAGCGATTTCGCCCGGGACACCCACTTCCTGCACCTGCGCAGCGCCA 1604

Qy   1537 CCCAGAACTAC------TTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGA 1590
Db   1605 GCCTCGGTTCCCAGCAGCTCTTGGGCCTGCCCCGAGACCCAGGGAGCAGCGTCAGCGGCA 1664

Qy   1591 CCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCC 1650
Db   1665 CCTTTGGCTGCCTGGGTGGGAGGCTCAGCATCCCCGGCACAGGGGTCAGCTTGCTGGTGC 1724

Qy   1651 CCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGG 1710
Db   1725 CCAATGGAGCCATTCCCCAGGGCAAGTTCTACGAGATGTATCTACTCATCAACAAGGCAG 1784

Qy   1711 AAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTG 1770
Db   1785 AAAGTACCCTGCCGCTTTCAGAAGGGACCCAGACAGTATTGAGCCCCTCGGTGACCTGTG 1844

Qy   1771 GACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGC 1830
Db   1845 GACCCACAGGCCTCCTGCTGTGCCGCCCCGTCATCCTCACCATGCCCCACTGTGCCGAAG 1904

Qy   1831 CCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGC 1890
Db   1905 TCAGTGCCCGTGACTGGATCTTTCAGCTCAAGACCCAGGCCCACCAGGGCCACTGGGAG- 1963

Qy   1891 AGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGG 1950
Db   1964 --GAGGTGGTGACCCTGGATGAGGAGACCCTGAACACACCCTGCTACTGCCAGCTGGAGC 2021

Qy   1951 CCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCC 2010
Db   2022 CCAGGGCCTGTCACATCCTGCTGGACCAGCTGGGCACCTACGTGTTCACGGGCGAGTCCT 2081

Qy   2011 TCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCT 2070
Db   2082 ATTCCCGCTCAGCAGTCAAGCGGCTCCAGCTGGCCGTCTTCGCCCCCGCCCTCTGCACCT 2141

Qy   2071 CCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGG 2130
Db   2142 CCCTGGAGTACAGCCTCCGGGTCTACTGCCTGGAGGACACGCCTGTAGCACTGAAGGAGG 2201

Qy   2131 TGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACT 2190
Db   2202 TGCTGGAGCTGGAGCGGACTCTGGGCGGATACTTGGTGGAGGAGCCGAAACCGCTAATGT 2261

Qy   2191 TCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGA 2250
Db   2262 TCAAGGACAGTTACCACAACCTGCGCCTCTCCCTCCATGACCTCCCCCATGCCCATTGGA 2321

Qy   2251 AGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGC 2310
Db   2322 GGAGCAAGCTGCTGGCCAAATACCAGGAGATCCCCTTCTATCACATTTGGAGTGGCAGCC 2381

Qy   2311 AGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGG 2370
Db   2382 AGAAGGCCCTCCACTGCACTTTCACCCTGGAGAGGCACAGCTTGGCCTCCACAGAGCTCA 2441

Qy   2371 CCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCA 2430
Db   2442 CCTGCAAGATCTGCGTGCGGCAAGTGGAAGGGGAGGGCCAGATATTCCAGCTGCATACCA 2501

Qy   2431 ACATCACCAAG---GACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGG 2487
Db   2502 CTCTGGCAGAGACACCTGCTGGCTCCCTGGACACTCTCTGCTCTGCCCCTGGCAGCACTG 2561

Qy   2488 TCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAA 2547
Db   2562 TCACCACCCAGCTGGGACCTTATGCCTTCAAGATCCCACTGTCCATCCGCCAGAAGATAT 2621

Qy   2548 TTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAAC 2607
Db   2622 GCAACAGCCTAGATGCCCCCAACTCACGGGGCAATGACTGGCGGATGTTAGCACAGAAGC 2681

Qy   2608 TCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCC 2667
Db   2682 TCTCTATGGACCGGTACCTGAATTACTTTGCCACCAAAGCGAGCCCCACGGGTGTGATCC 2741

Qy   2668 TCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAG 2727
Db   2742 TGGACCTCTGGGAAGCTCTGCAGCAGGACGATGGGGACCTCAACAGCCTGGCGAGTGCCT 2801

Qy   2728 TGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2786
Db   2802 TGGAGGAGATGGGCAAGAGTGAGATGCTGGTGGCTGTGGCCACCGACGGGGACTGCTGA 2860
```

```
RESULT 14
ABK92062
ID   ABK92062 standard; DNA; 2995 BP.
XX
AC   ABK92062;
XX
DT   14-AUG-2002  (first entry)
XX
DE   DNA encoding novel UNC5 receptor-like protein.
XX
KW   Cancer; metabolic disorder; skin disorder; infectious disease; anorexia;
KW   behavioral disorder; valve disease; endocrine disorder; heart disorder;
KW   blood disorder; anxiety disorder; brain disorder; inflammatory disorder;
KW   neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW   immune disorder; haematopoietic disorder; dyslipidaemia;
KW   metabolic disturbance; obesity; metabolic syndrome X; wasting disorder;
KW   gene therapy; transgenic animal; human; gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO200229058-A2.
XX
PD   11-APR-2002.
XX
PF   05-OCT-2001; 2001WO-US031248.
XX
PR   05-OCT-2000; 2000US-0238323P.
PR   05-OCT-2000; 2000US-0238325P.
PR   06-OCT-2000; 2000US-0238372P.
PR   06-OCT-2000; 2000US-0238373P.
PR   06-OCT-2000; 2000US-0238379P.
PR   06-OCT-2000; 2000US-0238382P.
PR   06-OCT-2000; 2000US-0238383P.
PR   06-OCT-2000; 2000US-0238384P.
```

```
PR   06-OCT-2000; 2000US-0238397P.
PR   06-OCT-2000; 2000US-0238400P.
PR   06-OCT-2000; 2000US-0238401P.
PR   06-OCT-2000; 2000US-0238402P.
PR   14-MAR-2001; 2001US-0275892P.
PR   08-JUN-2001; 2001US-0296860P.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Taupier RJ,  Burgess CE,  Zerhusen BD,  Mezes PS;
PI   Rastelli L,  Malyankar UM,  Grosse WM,  Alsobrook JP,  Lepley DM;
PI   Spytek KA,  Li L,  Edinger S,  Gerlach V,  Ellerman K,  Macdougall J;
PI   Gunther E,  Millet I,  Stone D,  Smithson G,  Szekeres ES;
XX
DR   WPI; 2002-444103/47.
DR   P-PSDB; ABG61795.
XX
PT   Novel isolated polypeptide, designated NOVX, useful for treating or
PT   preventing cancer, diabetes, obesity, dyslipidemia, anorexia, and
PT   metabolic, neurodegenerative, immune and hematopoietic disorders.
XX
PS   Claim 9; Page 93; 316pp; English.
XX
CC   The invention describes an isolated polypeptide (I), useful in the
CC   manufacture of a medicament for treating a syndrome associated with a
CC   human disease. (I), the polynucleotide encoding it (II) and an antibody
CC   (III) to (I) are useful for treating or preventing cancer, metabolic
CC   disorders, skin disorders, infectious disease, anorexia, behavioral
CC   disorders, valve diseases, endocrine disorders, heart and blood
CC   disorders, anxiety disorders, brain disorders, inflammatory disorders,
CC   neurodegenerative disorder, Alzheimer's disease, Parkinson's disorder,
CC   immune disorders, haematopoietic disorders, and the various
CC   dyslipidaemias, metabolic disturbances associated with obesity, and the
CC   metabolic syndrome X and wasting disorders associated with chronic
CC   diseases. (I), (II) or (III) are useful in screening assays, detection
CC   assays (e.g., chromosomal mapping, tissue typing, forensic biology),
CC   predictive medicine (e.g., diagnostic assays and prognostic assays), and
CC   in methods of treatment (e.g., therapeutic and prophylactic). (I) is
CC   useful as immunogen to produce antibodies immunospecific for (I), to
CC   screen for potential agonist and antagonist compounds, and as bait
CC   protein in a two-hybrid or three-hybrid assay. (II) is useful in gene
CC   therapy, to express (I), to detect (I) mRNA or a genetic lesion in a (I)
CC   gene, and to modulate activity of (I). A cell containing a vector
CC   expressing (I) is useful for producing non-human transgenic animals. This
CC   sequence encodes a novel human polypeptide described in the invention
XX
SQ   Sequence 2995 BP; 571 A; 984 C; 925 G; 515 T; 0 U; 0 Other;

  Query Match             30.4%;  Score 875.8;  DB 6;  Length 2995;
  Best Local Similarity   59.8%;  Pred. No. 5.3e-152;
  Matches 1727;  Conservative    0;  Mismatches 1022;  Indels  138;  Gaps   10;

Qy         38 CCGCCCGCGGGGCCCCGCGCCCGGCCCGCCCGCCTGCCCGCCCGCGGCCATGGCCGTCCG 97

Db         24 CAGCCCTGGGGGAGAGGCGCCCGAACCAGGCCGCGGGAGCATGGGGGCCCGGAGCGGAGC 83

Qy         98 GCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGG 157

Db         84 TCGGGGCGCGCTGCTGCTGGCACTGCTGCTCTGCTGGGACCCGAGGCTGAGCCAAGCAGG 143

Qy        158 TGCCCAGC---AGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCT 214

Db        144 TAGGAAGCGATCGGGTGAGGTGCTCCCTGACTCCTTCCCGTCAGCGCCAGCAGAGCCGCT 203

Qy        215 TCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCT 274

Db        204 GCCCTACTTCCTGCAGGAGCCACAGGACGCCTACATTGTGAAGAACAAGCCTGTGGAGCT 263

Qy        275 TGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCG 334

Db        264 CCGCTGCCGCGCCTTCCCCGCCACACAGATCTACTTCAAGTGCAACGGCGAGTGGGTCAG 323

Qy        335 CCAGGTGGACCACG---------------TGATCGAGCGCAGCACAGACGGGAGCAGTGG 379

Db        324 CCAGAACGACCACGTCACACAGGAAGGCCTGGATGAGGCCACCCTGGGGGCGCGGGGCGG 383

Qy        380 TGAGCCGACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGG 439

Db        384 CCTGCGGGTGCGCGAGGTGCAGATCGAGGTGTCGCGGCAGCAGGTGGAGGAGCTCTTTGG 443

Qy        440 GCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCA 499

Db        444 GCTGGAGGATTACTGGTGCCAGTGCGTGGCCTGGAGCTCCGCGGGCACCACCAAGAGTCG 503

Qy        500 GAAGGCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAA 559

Db        504 CCGAGCCTACGTCCGCATCGCCTGTCTGCGCAAGAACTTCGATCAGGAGCCTCTGGGCAA 563

Qy        560 GGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCC 619

Db        564 GGAGGTGCCCCTGGACCATGAGGTTCTCCTGCAGTGCCGCCCGCCGGAGGGGGTGCCTGT 623

Qy        620 AGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGT 679

Db        624 GGCCGAGGTGGAATGGCTCAAGAATGAGGATGTCATCGACCCCACCCAGGACACCAACTT 683

Qy        680 ATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAA 739

Db        684 CCTGCTCACCATCGACCACAACCTCATCATCCGCCAGGCCCGCCTGTCGGACACTGCCAA 743

Qy        740 CTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCAT 799

Db        744 CTATACCTGCGTGGCCAAGAACATCGTGGCCAAACGCCGGAGCACCACTGCCACCGTCAT 803

Qy        800 CGTCTACGTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTG 859

Db        804 CGTCTACGTGAATGGCGGCTGGTCCAGCTGGGCAGAGTGGTCACCCTGCTCCAACCGCTG 863

Qy        860 TGGGCGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGG 919

Db        864 TGGCCGAGGCTGGCAGAAGCGCACCCGGACCTGCACCAACCCCGCTCCACTCAACGGAGG 923

Qy        920 CGCTTTCTGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGT 979

Db        924 GGCCTTCTGCGAGGGCCAGGCATTCCAGAA---GACCGCCTGCACCACCATCTGCCCAGT 980

Qy        980 GGACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTG 1039

Db        981 CGATGGGGCGTGGACGGAGTGGAGCAAGTGGTCAGCCTGCAGCACTGAGTGTGCCCACTG 1040

Qy       1040 GCGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCAC 1099

Db       1041 GCGTAGCCGCGAGTGCATGGCGCCCCCACCCCAGAACGGAGGCCGTGACTGCAGCGGGAC 1100

Qy       1100 TGACCTGGACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCC---- 1155

Db       1101 GCTGCTCGACTCTAAGAACTGCACAGATGGGCTGTGCATGCAAAGTGAGCCTGTCCCCGC 1160

Qy       1156 --------------CTGAGGACGTGGCCCTCTATGTGGGCCTC---ATCGCCGTGGCCGT 1198

Db       1161 AGTGCTGGAGGCCTCAGGGGATGCGGCGCTGTATGCGGGGCTCGTGGTGGCCATCTTCGT 1220

Qy       1199 CTGCCTGGTCCTGCTGCTGCTTGTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCT 1258

Db       1221 GGTCGTGGCAATCCTCATGGCGGTGGGGGTGGTGGTGTACCGCCGCAACTGCCGTGACTT 1280

Qy       1259 GGACTCAGATGTGGCTGACTCGTCCATT---CTCACCTCAGGCTTCCAGCCCGTCAGCAT 1315

Db       1281 CGACACAGACATCACTGACTCATCTGCTGCCCTGACTGGTGGTTTCCACCCCGTCAACTT 1340

Qy       1316 CAAGCCCAGCAAAGCAGACAACCCCCATCTGCT------CACCATCCAGCCGGACCTCAG 1369

Db       1341 TAAGACGGCAAGGCCCAGTAACCCGCAGCTCCTACACCCCTCTGTGCCTCCTGACCTGAC 1400

Qy       1370 CACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCCAGCCCCAAGTT 1429
```

```
Db     1401 AGCCAGCGCCGGCATCTACCGCGGACCCGTGTATGCCCTGCAGGACTCCACCGACAAAAT 1460

Qy     1430 CCAGCTCACCAATGGGCACCTGCTCAGCCCC----------------------------- 1460

Db     1461 CCCCATGACCAACTCTCCTCTGCTGGACCCCTTACCCAGCCTTAAGGTCAAGGTCTACAG 1520

Qy     1461 -------------------------------------------------CTGGGTGGCGG 1471

Db     1521 CTCCAGCACCACGGGCTCTGGGCCAGGCCTGGCAGATGGGGCTGACCTGCTGGGGGTCTT 1580

Qy     1472 CCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCT 1531

Db     1581 GCCGCCTGGCACATACCCTAGCGATTTCGCCCGGGACACCCACTTCCTGCACCTGCGCAG 1640

Qy     1532 CTCCACCC------AGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTA 1585

Db     1641 CGCCAGCCTCGGTTCCCAGCAGCTCTTGGGCCTGCCCCGAGACCCAGGGAGCAGCGTCAG 1700

Qy     1586 TGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCT 1645

Db     1701 CGGCACCTTTGGCTGCCTGGGTGGGAGGCTCAGCATCCCCGGCACAGGTGTCAGCTTGCT 1760

Qy     1646 CATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAA 1705

Db     1761 GGTGCCCAATGGAGCCATTCCCCAGGGCAAGTTCTACGAGATGTATCTACTCATCAACAA 1820

Qy     1706 GCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAG 1765

Db     1821 GGCAGAAAGTACCCTGCCGCTTTCAGAAGGGACCCAGACAGTATTGAGCCCCTCGGTGAC 1880

Qy     1766 CTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGG 1825

Db     1881 CTGTGGACCCACAGGCCTCCTGCTGTGCCGCCCCGTCATCCTCACCATGCCCCACTGTGC 1940

Qy     1826 GGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTG 1885

Db     1941 CGAAGTCAGTGCCCGTGACTGGATCTTTCAGCTCAAGACCCAGGCCCACCAGGGCCACTG 2000

Qy     1886 GGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCT 1945

Db     2001 GGAGCAGGAGGTGGTGACCCTGGATGAGGAGACCCTGAACACACCCTGCTACTGCCAGCT 2060

Qy     1946 GGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGA 2005

Db     2061 GGAGCCCAGGGCCTGTCACATCCTGCTGGACCAGCTGGGCACCTACGTGTTCACGGGCGA 2120

Qy     2006 GGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTG 2065

Db     2121 GTCCTATTCCCGCTCAGCAGTCAAGCGGCTCCAGCTGGCCGTCTTCGCCCCCGCCCTCTG 2180

Qy     2066 CACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAA 2125

Db     2181 CACCTCCCTGGAGTACAGCCTCCGGGTCTACTGCCTGGAGGACACGCCTGTAGCACTGAA 2240

Qy     2126 GGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCT 2185

Db     2241 GGAGGTGCTGGAGCTGGAGCGGACTCTGGGCGGATACTTGGTGGAGGAGCCGAAACCGCT 2300

Qy     2186 GCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCT 2245

Db     2301 AATGTTCAAGGACAGTTACCACAACCTGCGCCTCTCCCTCCATGACCTCCCCCATGCCCA 2360

Qy     2246 GTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGG 2305

Db     2361 TTGGAGGAGCAAGCTGCTGGCCAAATACCAGGAGATCCCCTTCTATCACATTTGGAGTGG 2420

Qy     2306 CACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGA 2365

Db     2421 CAGCCAGAAGGCCCTCCACTGCACTTTCACCCTGGAGAGGCACAGCTTGGCCTCCACAGA 2480

Qy     2366 CCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAA 2425

Db     2481 GCTCACCTGCAAGATCTGCGTGCGGCAAGTGGAAGGGGAGGGCCAGATATTCCAGCTGCA 2540

Qy     2426 CTTCAACATCACCAAGGACA---CAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGC 2482

Db     2541 TACCACTCTGGCAGAGACACCTGCTGGCTCCCTGGACACTCTCTGCTCTGCCCCTGGCAG 2600

Qy     2483 GGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAA 2542

Db     2601 CACTGTCACCACCCAGCTGGGACCTTATGCCTTCAAGATCCCACTGTCCATCCGCCAGAA 2660

Qy     2543 GATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCA 2602

Db     2661 GATATGCAACAGCCTAGATGCCCCCAACTCACGGGGCAATGACTGGCGGATGTTAGCACA 2720

Qy     2603 GAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCAT 2662

Db     2721 GAAGCTCTCTATGGACCGGTACCTGAATTACTTTGCCACCAAAGCGAGCCCCACGGGTGT 2780

Qy     2663 GATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGC 2722

Db     2781 GATCCTGGACCTCTGGGAAGCTCTGCAGCAGGACGATGGGGACCTCAACAGCCTGGCGAG 2840

Qy     2723 AGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTG 2782

Db     2841 TGCCTTGGAGGAGATGGGCAAGAGTGAGATGCTGGTGGCTGTGGCCACCGACGGGGACTG 2900

Qy     2783 CTGAGGC 2789

Db     2901 CTGAGCC 2907
```

```
RESULT 15
ABQ93898
ID   ABQ93898 standard; DNA; 2895 BP.
XX
AC   ABQ93898;
XX
DT   01-NOV-2002  (first entry)
XX
DE   Human transmembrane receptor UNC5H2-like NOV11 DNA, SEQ ID NO:37.
XX
KW   Human; NOVX; neurological disorder; Alzheimer's disease;
KW   Huntington's disease; Parkinson's disease; pain; behavioural disorder;
KW   addiction; tuberous sclerosis; cancer; immune disorder; allergy;
KW   autoimmune disease; myasthenia gravis; asthma; arthritis; diabetes;
KW   thyroiditis; cardiovascular disease; hypertension; reproductive disorder;
KW   endometriosis; incontinence; psoriasis; scleroderma; alopecia; ulcer;
KW   pancreatitis; cirrhosis; glomerular endotheliosis; bacterial infection;
KW   polycystic kidney disease; endocrine disorder; obesity; cardiomyopathy;
KW   atherosclerosis; cell signal processing-related disorder;
KW   metabolic pathway regulation disorder; cytostatic; neuroprotective;
KW   antiinflammatory; immunosuppressive; analgesic; antiatherosclerotic;
KW   dermatological; antibacterial; antiarthritic; hepatotropic; neurogenesis;
KW   differentiation; proliferation; motility; haematopoiesis; wound healing;
KW   angiogenesis; forensic biology; transgenic animal; drug screening;
KW   gene therapy; NOV11; transmembrane receptor UNC5H2-like; chromosome 10;
KW   gene; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   5'UTR           1..30
FT                   /*tag=  a
FT   CDS             31..2868
FT                   /*tag=  b
FT                   /product= "NOV11"
FT   sig_peptide     31..108
FT                   /*tag=  c
FT   mat_peptide     109..2865
FT                   /*tag=  d
FT                   /product= "Mature NOV11 protein"
FT   3'UTR           2869..2895
FT                   /*tag=  e
XX
```

```
PN   WO200253742-A2.
XX
PD   11-JUL-2002.
XX
PF   07-JAN-2002; 2002WO-US000375.
XX
PR   05-JAN-2001; 2001US-0260018P.
PR   08-JAN-2001; 2001US-0260360P.
PR   28-FEB-2001; 2001US-0272411P.
PR   02-MAR-2001; 2001US-0272817P.
PR   05-JUL-2001; 2001US-0303231P.
PR   12-JUL-2001; 2001US-0305060P.
PR   10-SEP-2001; 2001US-0318405P.
PR   12-SEP-2001; 2001US-0318700P.
PR   04-JAN-2002; 2002US-00037417.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Kekuda R,  Alsobrook JP,  Tchernev VT,  Liu X,  Spytek KA;
PI   Patturajan M,  Grosse WM,  Lepley DM,  Burgess CE,  Vernet CAM,  Li L;
PI   Gorman L,  Edinger S,  Sciore P,  Ellerman K,  Malyankar U;
PI   Rothenberg M,  Stone D,  Boldog F,  Guo X,  Shenoy S,  Anderson D;
PI   Padigaru M,  Taupier RJ,  Miller CE,  Eisen A;
XX
DR   WPI; 2002-583619/62.
DR   P-PSDB; ABB09520.
XX
PT   Novel polypeptides and nucleic acids homologous to transmembrane
PT   receptor, thymosin, neuromodulin-like family of proteins for diagnosing,
PT   treating cancer, atherosclerosis, neurological, skin and autoimmune
PT   disorders.
XX
PS   Claim 9a; Page 121; 323pp; English.
XX
CC   The invention relates to 24 novel human proteins designated NOV1-NOV14
CC   (ABB09501-ABB09524), collectively referred to as NOVX proteins, and
CC   nucleic acids encoding them (ABQ93879-ABQ93902). NOVX proteins and
CC   nucleotides are useful in the treatment, diagnosis or prevention of NOVX-
CC   associated disorders or in the manufacture of a medicament for treating
CC   such disorders, with specific applications described for each of the 24
CC   NOVX proteins, based on their homology to known proteins. Various
CC   disorders are associated with NOVX proteins including neurological
CC   disorders (e.g., Alzheimer's, Huntington's and Parkinson's diseases),
CC   pain, behavioural disorders, addiction, tuberous sclerosis, cancers
CC   (e.g., colorectal cancer, leukaemia and osteosarcoma), immune disorders
CC   (e.g., allergies and autoimmune diseases), myasthenia gravis, asthma,
CC   various forms of arthritis, diabetes, thyroiditis, cardiovascular disease
CC   (e.g., hypertension), reproductive disorders, endometriosis,
CC   incontinence, psoriasis, scleroderma, alopecia, ulcers, pancreatitis,
CC   cirrhosis, glomerular endotheliosis, polycystic kidney disease, endocrine
CC   disorders, obesity, bacterial infections and particularly cardiomyopathy,
CC   atherosclerosis, cell signal processing-related disorders and disorders
CC   of metabolic pathway regulation. NOVX nucleic acids and polypeptides may
CC   be used to identify cellular receptors or downstream effectors which
CC   binds to a NOVX protein, and are also useful as targets for the
CC   identification of small molecules that modulate or inhibit processes such
CC   as neurogenesis, cell differentiation, cell motility, cellular
CC   proliferation, haematopoiesis, wound healing and angiogenesis. NOVX
CC   nucleic acid sequences can be used to identify a cell or tissue type and
CC   are useful as a source of primers or probes for forensic biology and for
CC   identifying and cloning NOVX homologues in other cell types. Cells
CC   comprising NOVX nucleic acids are useful for producing non-human
CC   transgenic animals which are useful for studying the function and
CC   activity of NOVX proteins and for identifying and evaluating modulators
CC   of NOVX activity. The present sequence represents DNA encoding the
CC   transmembrane receptor UNC5H2-like protein NOV11. The gene encoding NOV11
CC   is located on chromosome 10
XX
SQ   Sequence 2895 BP; 557 A; 960 C; 854 G; 524 T; 0 U; 0 Other;

  Query Match             30.1%;  Score 866.4;  DB 6;  Length 2895;
  Best Local Similarity   60.6%;  Pred. No. 2.8e-150;
  Matches 1659;  Conservative    0;  Mismatches  941;  Indels  138;  Gaps    9;

Qy        184 ACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATG 243
Db        140 ACTCCTTCCCGTCAGCGCCAGCAGAGCCGCTGCCCTACTTCCTGCAGGAGCCACAGGACG 199

Qy        244 TGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGA 303
Db        200 CCTACATTGTGAAGAACAAGCCTGTGGAGCTCCGCTGCCGCGCCTTCCCCGCCACACAGA 259

Qy        304 TCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCA 363
Db        260 TCTACTTCAAGTGCAACGGCGAGTGGGTCAGCCAGAACGACCACGTCACACAGGAAGGCC 319

Qy        364 CAGACGGGAGCAGTGGTGAGCCGACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGG 423
Db        320 TGGATGAGGCCACCGGTCTGCGGGTGCGCGAGGTGCAGATCGAGGTGTCGCGGCAGCAGG 379

Qy        424 TCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGG 483
Db        380 TGGAGGAGCTCTTTGGGCTGGAGGATTACTGGTGCCAGTGCGTGGCCTGGAGCTCCGCGG 439

Qy        484 GCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGC 543
Db        440 GCACCACCAAGAGTCGCCGAGCCTACGTCCGCATCGCCTGTCTGCGCAAGAACTTCGATC 499

Qy        544 AGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCAC 603
Db        500 AGGAGCCTCTGGGCAAGGAGGTGCCCCTGGACCATGAGGTTCTCCTGCAGTGCCGCCCGC 559

Qy        604 CGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGT 663
Db        560 CGGAGGGGGTGCCTGTGGCCGAGGTGGAATGGCTCAAGAATGAGGATGTCATCGACCCCA 619

Qy        664 CCCTGGACCCCAATGTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCC 723
Db        620 CCCAGGACACCAACTTCCTGCTCACCATCGACCACAACCTCATCATCCGCCAGGCCCGCC 679

Qy        724 TTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCG 783
Db        680 TGTCGGACACTGCCAACTATACCTGCGTGGCCAAGAACATCGTGGCCAAACGCCGGAGCA 739

Qy        784 CCTCCGCTGCTGTCATCGTCTACGTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCG 843
Db        740 CCACTGCCACCGTCATCGTCTACGTGAATGGCGGCTGGTCCAGCTGGGCAGAGTGGTCAC 799

Qy        844 TCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGG 903
Db        800 CCTGCTCCAACCGCTGTGGCCGAGGCTGGCAGAAGCGCACCCGGACCTGCACCAACCCCG 859

Qy        904 CGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCCT 963
Db        860 CTCCACTCAACGGAGGGGCCTTCTGCGAGGGCCAGGCATTCCAGAA---GACCGCCTGCA 916

Qy        964 CTCTGCTTGTCTCTGTGGACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGC 1023
Db        917 CCACCATCTGCCCAGTCGATGGGGCGTGGACGGAGTGGAGCAAGTGGTCAGCCTGCAGCA 976

Qy       1024 TGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGG 1083
Db        977 CTGAGTGTGCCCACTGGCGTAGCCGCGAGTGCATGGCGCCCCCACCCCAGAACGGAGGCC 1036

Qy       1084 AGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGTACCAGTGACCTCTGTGTACACA 1143
Db       1037 GTGACTGCAGCGGGACGCTGCTCGACTCTAAGAACTGCACAGATGGGCTGTGCATGCAAA 1096

Qy       1144 GTGCTTC-------------------------------TGGCCCTGAGGACGTGGCCC 1170
Db       1097 GTGAGTCACAGTGTGGTCCTCCTGTCCCCGCAGTGCTGGAGGCCTCAGGGGATGCGGCGC 1156

Qy       1171 TCTATGTGGGCCTC---ATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCA 1227
Db       1157 TGTATGCGGGGCTCGTGGTGGCCATCTTCGTGGTCGTGGCAATCCTCATGGCGGTGGGGG 1216
```

```
Qy    1228 TCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT- 1286
Db    1217 TGGTGGTGTACCGCCGCAACTGCCGTGACTTCGACACAGACATCACTGACTCATCTGCTG 1276

Qy    1287 --CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATC 1344
Db    1277 CCCTGACTGGTGGTTTCCACCCCGTCAACTTTAAGACGGCAAGGCCCAGTAACCCGCAGC 1336

Qy    1345 TGCT------CACCATCCAGCCGGACCTCAGCACCACCACCACCTACCAGGGCAGTCTCT 1398
Db    1337 TCCTACACCCCTCTGTGCCTCCTGACCTGACAGCCAGCGCCGGCATCTACCGCGGACCCG 1396

Qy    1399 GTCCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCC 1458
Db    1397 TGTATGCCCTGCAGGACTCCACCGACAAAATCCCCATGACCAACTCTCCTCTGCTGGACC 1456

Qy    1459 CC---------------------------------------------------------- 1460
Db    1457 CCTTACCCAGCCTTAAGGTCAAGGTCTACAGCTCCAGCACCACGGGCTCTGGGCCAGGCC 1516

Qy    1461 --------------------CTGGGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCA 1500
Db    1517 TGGCAGATGGGGCTGACCTGCTGGGGGTCTTGCCGCCTGGCACATACCCTAGCGATTTCG 1576

Qy    1501 CCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACCC------AGAACTACTTCCGCT 1554
Db    1577 CCCGGGACACCCACTTCCTGCACCTGCGCAGCGCCAGCCTCGGTTCCCAGCAGCTCTTGG 1636

Qy    1555 CCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGC 1614
Db    1637 GCCTGCCCCGAGACCCAGGGAGCAGCGTCAGCGGCACCTTTGGCTGCCTGGGTGGAGGC  1696

Qy    1615 TGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCCAGATGCCATACCCCGAGGGA 1674
Db    1697 TCAGCATCCCCGGCACAGGTGTCAGCTTGCTGGTGCCCAATGGAGCCATTCCCCAGGGCA 1756

Qy    1675 AGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTG 1734
Db    1757 AGTTCTACGAGATGTATCTACTCATCAACAAGGCAGAAAGTACCCTGCCGCTTTCAGAAG 1816

Qy    1735 GCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCTGGCGTCCTGCTCACCC 1794
Db    1817 GGACCCAGACAGTATTGAGCCCCTCGGTGACCTGTGGACCCACAGGCCTCCTGCTGTGCC 1876

Qy    1795 GGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGC 1854
Db    1877 GCCCCGTCATCCTCACCATGCCCCACTGTGCCGAAGTCAGTGCCCGTGACTGGATCTTTC 1936

Qy    1855 GCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGCAGGATGTGCTGCACCTGGGCGAGG 1914
Db    1937 AGCTCAAGACCCAGGCCCACCAGGGCCACTGGG---AGGAGGTGGTGACCCTGGATGAGG 1993

Qy    1915 AGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCG 1974
Db    1994 AGACCCTGAACACACCCTGCTACTGCCAGCTGGAGCCCAGGGCCTGTCACATCCTGCTGG 2053

Qy    1975 AGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCC 2034
Db    2054 ACCAGCTGGGCACCTACGTGTTCACGGGCGAGTCCTATTCCCGCTCAGCAGTCAAGCGGC 2113

Qy    2035 TCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCT 2094
Db    2114 TCCAGCTGGCCGTCTTCGCCCCCGCCCTCTGCACCTCCCTGGAGTACAGCCTCCGGGTCT 2173

Qy    2095 ACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGG 2154
Db    2174 ACTGCCTGGAGGACACGCCTGTAGCACTGAAGGAGGTGCTGGAGCTGGAGCGGACTCTGG 2233

Qy    2155 GGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAGTTACCACAACCTGC 2214
Db    2234 GCGGATACTTGGTGGAGGAGCCGAAACCGCTAATGTTCAAGGACAGTTACCACAACCTGC 2293

Qy    2215 GCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACC 2274
Db    2294 GCCTCTCCCTCCATGACCTCCCCCATGCCCATTGGAGGAGCAAGCTGCTGGCCAAATACC 2353

Qy    2275 AGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTTGCACTGCACCTTCA 2334
Db    2354 AGGAGATCCCCTTCTATCACATTTGGAGTGGCAGCCAGAAGGCCCTCCACTGCACTTTCA 2413

Qy    2335 CCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGG 2394
Db    2414 CCCTGGAGAGGCACAGCTTGGCCTCCACAGAGCTCACCTGCAAGATCTGCGTGCGGCAAG 2473

Qy    2395 TGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACCAAG---GACACAAGGT 2451
Db    2474 TGGAAGGGGAGGGCCAGATATTCCAGCTGCATACCACTCTGGCAGAGACACCTGCTGGCT 2533

Qy    2452 TTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTG 2511
Db    2534 CCCTGGACACTCTCTGCTCTGCCCCTGGCAGCACTGTCACCACCCAGCTGGGACCTTATG 2593

Qy    2512 CCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTGGACCCACCCTGTA 2571
Db    2594 CCTTCAAGATCCCACTGTCCATCCGCCAGAAGATATGCAACAGCCTAGATGCCCCCAACT 2653

Qy    2572 GGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGACAGCCATCTCAGCT 2631
Db    2654 CACGGGGCAATGACTGGCGGATGTTAGCACAGAAGCTCTCTATGGACCGGTACCTGAATT 2713

Qy    2632 TCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGGGAGGCGCGGCACT 2691
Db    2714 ACTTTGCCACCAAAGCGAGCCCCACGGGTGTGATCCTGGACCTCTGGGAAGCTCTGCAGC 2773

Qy    2692 TCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGGCCAGCCAGACG 2751
Db    2774 AGGACGATGGGGACCTCAACAGCCTGGCGAGTGCCTTGGAGGAGATGGGCAAGAGTGAGA 2833

Qy    2752 CTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGAGGC 2789
Db    2834 TGCTGGTGGCTGTGGCCACCGACGGGGACTGCTGAGCC 2871
```

Search completed: October 6, 2004, 12:44:56
Job time : 1102 secs

GenCore version 5.1.6


OM nucleic - nucleic search, using sw model

```
Run on:          October  6, 2004, 12:08:07 ; Search time 197 Seconds
                                              (without alignments)
                                              8115.805 Million cell updates/sec

Title:           US-09-970-944-1
Perfect score:   2881
Sequence:        1 agctggggctccgggctgag..........ccttccccacaccggggaga 2881

Scoring table:   IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched:        682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters:      1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Issued_Patents_NA:*
               1:  /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
               2:  /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
               3:  /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
               4:  /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
               5:  /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
               6:  /cgn2_6/ptodata/2/ina/backfiles1.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| | Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|---|
| | 1 | 2235.8 | 77.6 | 3014 | 2 | US-08-808-982-1 | Sequence 1, Appli |
| | 2 | 2235.8 | 77.6 | 3014 | 3 | US-09-306-902A-1 | Sequence 1, Appli |
| | 3 | 1570.4 | 54.5 | 1787 | 2 | US-08-808-982-2 | Sequence 2, Appli |
| | 4 | 1570.4 | 54.5 | 1787 | 3 | US-09-306-902A-2 | Sequence 2, Appli |
| | 5 | 1147.8 | 39.8 | 1282 | 4 | US-09-833-381-1806 | Sequence 1806, Ap |
| | 6 | 783.4 | 27.2 | 2831 | 2 | US-08-808-982-3 | Sequence 3, Appli |
| | 7 | 783.4 | 27.2 | 2831 | 3 | US-09-306-902A-3 | Sequence 3, Appli |
| | 8 | 419 | 14.5 | 1605 | 4 | US-09-833-381-1807 | Sequence 1807, Ap |
| c | 9 | 256 | 8.9 | 771 | 1 | US-08-253-155A-17 | Sequence 17, Appl |
| | 10 | 119 | 4.1 | 305 | 2 | US-08-808-982-4 | Sequence 4, Appli |
| | 11 | 119 | 4.1 | 305 | 3 | US-09-306-902A-4 | Sequence 4, Appli |
| | 12 | 51.8 | 1.8 | 1280 | 3 | US-08-483-533-38 | Sequence 38, Appl |
| | 13 | 51.8 | 1.8 | 1280 | 4 | US-09-283-471A-38 | Sequence 38, Appl |
| | 14 | 51.2 | 1.8 | 699 | 4 | US-09-252-991A-7947 | Sequence 7947, Ap |
| | 15 | 51.2 | 1.8 | 1302 | 4 | US-09-252-991A-7809 | Sequence 7809, Ap |
| c | 16 | 51.2 | 1.8 | 1947 | 4 | US-09-252-991A-7533 | Sequence 7533, Ap |
| | 17 | 50.8 | 1.8 | 1770 | 4 | US-09-252-991A-12265 | Sequence 12265, A |
| | 18 | 50.8 | 1.8 | 3885 | 4 | US-09-369-364A-16 | Sequence 16, Appl |
| | 19 | 50.8 | 1.8 | 4884 | 4 | US-09-252-991A-12126 | Sequence 12126, A |
| c | 20 | 50.8 | 1.8 | 4884 | 4 | US-09-252-991A-12292 | Sequence 12292, A |
| | 21 | 50.2 | 1.7 | 1300 | 3 | US-08-483-533-39 | Sequence 39, Appl |
| | 22 | 50.2 | 1.7 | 1300 | 4 | US-09-283-471A-39 | Sequence 39, Appl |
| c | 23 | 48.6 | 1.7 | 1068 | 4 | US-09-252-991A-9933 | Sequence 9933, Ap |
| | 24 | 48.6 | 1.7 | 1251 | 4 | US-09-252-991A-10142 | Sequence 10142, A |
| | 25 | 48 | 1.7 | 4257 | 2 | US-08-690-473-1 | Sequence 1, Appli |
| | 26 | 48 | 1.7 | 4257 | 3 | US-09-259-821A-1 | Sequence 1, Appli |
| | 27 | 48 | 1.7 | 4257 | 3 | US-08-843-659-1 | Sequence 1, Appli |
| c | 28 | 48 | 1.7 | 8312 | 4 | US-09-620-312D-1048 | Sequence 1048, Ap |
| | 29 | 47.8 | 1.7 | 1288 | 1 | US-08-440-856A-9 | Sequence 9, Appli |
| c | 30 | 47.8 | 1.7 | 2211 | 4 | US-09-252-991A-11256 | Sequence 11256, A |
| | 31 | 47.8 | 1.7 | 2289 | 4 | US-09-252-991A-10995 | Sequence 10995, A |
| | 32 | 47.6 | 1.7 | 1627 | 4 | US-09-675-018B-11 | Sequence 11, Appl |
| | 33 | 47 | 1.6 | 7218 | 1 | US-08-232-463-14 | Sequence 14, Appl |
| | 34 | 46.8 | 1.6 | 1951 | 3 | US-08-922-865-1 | Sequence 1, Appli |
| | 35 | 46.8 | 1.6 | 1951 | 4 | US-09-510-949-1 | Sequence 1, Appli |
| | 36 | 46.4 | 1.6 | 453 | 4 | US-09-252-991A-3971 | Sequence 3971, Ap |
| | 37 | 46.4 | 1.6 | 567 | 4 | US-09-252-991A-4004 | Sequence 4004, Ap |
| | 38 | 46.2 | 1.6 | 11220 | 3 | US-09-105-537-32 | Sequence 32, Appl |
| | 39 | 46.2 | 1.6 | 36778 | 3 | US-09-105-537-5 | Sequence 5, Appli |
| | 40 | 46.2 | 1.6 | 38506 | 3 | US-09-320-878-19 | Sequence 19, Appl |
| | 41 | 46.2 | 1.6 | 38506 | 4 | US-09-141-908-1 | Sequence 1, Appli |
| | 42 | 46.2 | 1.6 | 38506 | 4 | US-09-657-440-19 | Sequence 19, Appl |
| | 43 | 46 | 1.6 | 1065 | 4 | US-09-340-798A-31 | Sequence 31, Appl |
| | 44 | 46 | 1.6 | 3624 | 1 | US-07-951-715A-6 | Sequence 6, Appli |
| | 45 | 46 | 1.6 | 3624 | 2 | US-08-459-448A-6 | Sequence 6, Appli |

ALIGNMENTS

```
RESULT 1
US-08-808-982-1
; Sequence 1, Application US/08808982
; Patent No. 5939271
;  GENERAL INFORMATION:
;    APPLICANT:  Tessier-Lavigne, Marc
;    APPLICANT:  Leonardo, E. David
;    APPLICANT:  Hink, Lindsay
;    APPLICANT:  Masu, Masayuki
;    APPLICANT:  Kazuko, Keino-Masu
;    TITLE OF INVENTION:  Netrin Receptors
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SCIENCE & TECHNOLOGY LAW GROUP
;      STREET:  268 BUSH STREET, SUITE 3200
;      CITY:  SAN FRANCISCO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/808,982
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  OSMAN, RICHARD A
;      REGISTRATION NUMBER:  36,627
;      REFERENCE/DOCKET NUMBER:  UC96-217
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 343-4341
;      TELEFAX:  (415) 343-4342
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3014 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-808-982-1

  Query Match             77.6%;  Score 2235.8;  DB 2;  Length 3014;
  Best Local Similarity   88.5%;  Pred. No. 0;
  Matches 2474;  Conservative    0;  Mismatches  312;  Indels   11;  Gaps    4;

Qy         87 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 146
              |||||||||||||||||||||||||| |||||||||||||||||||||||| |||||
```

```
Db      1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGTGCTCCTGGGCATAGTCCTCGCCGCCTGGCTT 60

Qy    147 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 206
Db     61 CGTGGTTCGGGTGCCCAGCAGAGTGCCACGGTGGCCAATCCAGTGCCCGGTGCCAACCCC 120

Qy    207 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 266
Db    121 GACCTGCTGCCCCACTTCCTGGTAGAGCCTGAGGACGTGTACATTGTCAAGAACAAGCCG 180

Qy    267 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 326
Db    181 GTGTTGTTGGTGTGCAAGGCTGTGCCTGCCACCCAGATCTTCTTCAAGTGCAATGGGGAA 240

Qy    327 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCG 386
Db    241 TGGGTCCGCCAGGTCGATCACGTAATTGAACGCAGCACCGACAGCAGCAGCGGATTGCCA 300

Qy    387 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 446
Db    301 ACCATGGAGGTCCGTATCAACGTATCGAGGCAGCAGGTAGAGAAAGTGTTTGGGCTGGAG 360

Qy    447 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 506
Db    361 GAATACTGGTGCCAGTGTGTGGCATGGAGCTCCTCGGGTACCACCAAAAGTCAGAAGGCC 420

Qy    507 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 566
Db    421 TACATCCGGATTGCCTATTTGCGCAAGAACTTTGAGCAGGAGCCACTGGCCAAGGAAGTG 480

Qy    567 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 626
Db    481 TCACTGGAGCAAGGCATTGTACTACCTTGTCGCCCCCCAGAAGGAATCCCCCCAGCTGAG 540

Qy    627 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 686
Db    541 GTGGAGTGGCTTCGAAATGAGGACCTCGTGGACCCCTCCCTCGATCCCAATGTGTACATC 600

Qy    687 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 746
Db    601 ACGCGGGAGCACAGCCTAGTCGTGCGTCAGGCCCGCCTGGCCGACACGGCCAACTACACC 660

Qy    747 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 806
Db    661 TGTGTGGCCAAGAACATCGTAGCCCGTCGCCGAAGCACCTCTGCAGCGGTCATTGTTTAT 720

Qy    807 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 866
Db    721 GTGAACGGTGGGTGGTCGACGTGGACTGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGT 780

Qy    867 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 926
Db    781 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCACCTCTCAACGGGGGCGCCTTC 840

Qy    927 TGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGC 986
Db    841 TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACTCTG---TGCCCAGTGGATGGG 897

Qy    987 AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC 1046
Db    898 AGCTGGAGTTCGTGGAGTAAGTGGTCAGCCTGTGGGCTTGACTGCACCCACTGGCGGAGC 957

Qy   1047 CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG 1106
Db    958 CGCGAGTGCTCTGACCCAGCACCCCGCAATGGAGGTGAGGAGTGTCGGGGTGCTGACCTG 1017

Qy   1107 GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG 1166
Db   1018 GACACCCGCAACTGTACCAGTGACCTCTGCCTGCACACCGCTTCTTGCCCCGAGGACGTG 1077

Qy   1167 GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC 1226
Db   1078 GCTCTCTACATCGGCCTTGTCGCTGTGGCTGTGTGCCTCTTCTTGCTGTTGCTGGCCCTT 1137

Qy   1227 ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1286
Db   1138 GGACTCATTTACTGTCGCAAGAAGGAAGGGCTGGACTCCGATGTGGCCGACTCGTCCATC 1197

Qy   1287 CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1346
Db   1198 CTCACCTCGGGCTTCCAGCCTGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCACCTG 1257

Qy   1347 CTCACCATCCAGCCGGACCTCAGCACCACC---ACCACCTACCAGGGCAGTCTCTGTCCC 1403
Db   1258 CTCACCATCCAGCCAGACCTCAGCACCACCACTACCACCTACCAGGGCAGTCTATGTTCG 1317

Qy   1404 CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG 1463
Db   1318 AGGCAGGATGGACCCAGCCCCAAGTTCCAGCTCTCTAATGGTCACCTGCTCAGCCCACTG 1377

Qy   1464 GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC 1523
Db   1378 GGGAGTGGCCGCCATACGTTGCACCACAGCTCACCCACCTCTGAGGCTGAGGACTTCGTC 1437

Qy   1524 TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC 1583
Db   1438 TCCCGCCTCTCCACCCAAAACTACTTTCGTTCCCTGCCCCGCGGCACCAGCAACATGGCC 1497

Qy   1584 TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC 1643
Db   1498 TACGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACGGGGATCAGCCTC 1557

Qy   1644 CTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC 1703
Db   1558 CTCATACCCCCGGATGCCATCCCCCGAGGAAAGATCTACGAGATCTACCTCACACTGCAC 1617

Qy   1704 AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT 1763
Db   1618 AAGCCAGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCAGTCGTT 1677

Qy   1764 AGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT 1823
Db   1678 AGCTGTGGGCCCCCAGGAGTCCTGCTCACCCGGCCAGTCATCCTTGCAATGGACCACTGT 1737

Qy   1824 GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC 1883
Db   1738 GGAGAGCCCAGCCCTGACAGCTGGAGTCTGCGCCTCAAAAAGCAGTCCTGCGAGGGCAGT 1797

Qy   1884 TGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG 1943
Db   1798 TGGG---AGGATGTGCTGCACCTTGGTGAGGAGTCACCTTCCCACCTCTACTACTGCCAG 1854

Qy   1944 CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA 2003
Db   1855 CTGGAGGCCGGGGCCTGCTATGTCTTCACGGAGCAGCTGGGCCGCTTTGCCCTGGTAGGA 1914

Qy   2004 GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC 2063
Db   1915 GAGGCCCTCAGCGTGGCTGCCACCAAGCGCCTCAGGCTCCTTCTGTTTGCTCCCGTGGCC 1974

Qy   2064 TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC 2123
Db   1975 TGTACGTCCCTTGAGTACAACATCCGAGTGTACTGCCTACACGACACCCACGACGCTCTC 2034

Qy   2124 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTC 2183
Db   2035 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTAGGTGGACAGCTGATCCAGGAGCCTCGCGTC 2094

Qy   2184 CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCC 2243
Db   2095 CTGCACTTCAAAGACAGTTACCACAACCTACGTCTCTCCATCCACGACGTGCCCAGCTCC 2154

Qy   2244 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2303
Db   2155 CTGTGGAAGAGCAAGCTACTTGTCAGCTACCAGGAGATCCCTTTTTACCACATCTGGAAC 2214
```

```
Qy       2304 GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2363

Db       2215 GGCACCCAGCAGTATCTGCACTGCACCTTCACCCTGGAGCGCATCAACGCCAGCACCAGC 2274

Qy       2364 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2423

Db       2275 GACCTGGCCTGCAAGGTGTGGGTGTGGCAGGTGGAGGGAGATGGGCAGAGCTTCAACATC 2334

Qy       2424 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2483

Db       2335 AACTTCAACATCACTAAGGACACAAGGTTTGCTGAATTGTTGGCTCTGGAGAGTGAAGGG 2394

Qy       2484 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2543

Db       2395 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAAAAG 2454

Qy       2544 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2603

Db       2455 ATCATCGCCAGTCTGGACCCACCCTGCAGCCGGGGCGCCGACTGGAGAACTCTAGCCCAG 2514

Qy       2604 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG 2663

Db       2515 AAACTTCACCTGGACAGCCATCTTAGCTTCTTTGCCTCCAAGCCCAGCCCTACAGCCATG 2574

Qy       2664 ATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2723

Db       2575 ATCCTCAACCTATGGGAGGCACGGCACTTCCCCAACGGCAACCTCGGCCAGCTGGCAGCA 2634

Qy       2724 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2783

Db       2635 GCTGTGGCCGGACTGGGCCAACCAGATGCTGGCCTCTTCACGGTGTCGGAGGCCGAGTGT 2694

Qy       2784 TGAGGCCGGCCAGGCCCG--ACACCTACACTCTCACCAGCTTTGGCACCCACCAAGGACA 2841

Db       2695 TGAGACCAGCCAGGCCGGTAATGCCTACATTCTCACCAGCTTTGACACCTGCCAGGGACA 2754

Qy       2842 GGCAGAAGCCGGACAGGGGCCCTTCCCCACACCGGGG 2878

Db       2755 GGCAAAACCAGACAGGGGCCCTTCCCCCACACCCGGG 2791
```

```
RESULT 2
US-09-306-902A-1
; Sequence 1, Application US/09306902A
; Patent No. 6277585
;  GENERAL INFORMATION:
;       APPLICANT: Tessier-Lavigne, Marc
;                  Leonardo, E. David
;                  Hink, Lindsay
;                  Masu, Masayuki
;                  Kazuko, Keino-Masu
;       TITLE OF INVENTION: Netrin Receptors
;       NUMBER OF SEQUENCES: 9
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
;            STREET: 268 BUSH STREET, SUITE 3200
;            CITY: SAN FRANCISCO
;            STATE: CALIFORNIA
;            COUNTRY: USA
;            ZIP: 94104
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/306,902A
;            FILING DATE: 07-May-1999
;            CLASSIFICATION: &lt;Unknown&gt;
;       ATTORNEY/AGENT INFORMATION:
;            NAME: OSMAN, RICHARD A
;            REGISTRATION NUMBER: 36,627
;            REFERENCE/DOCKET NUMBER: UC96-217
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (415) 343-4341
;            TELEFAX: (415) 343-4342
;  INFORMATION FOR SEQ ID NO: 1:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 3014 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: double
;            TOPOLOGY: linear
;       MOLECULE TYPE: cDNA
;       SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-306-902A-1

  Query Match             77.6%;  Score 2235.8;  DB 3;  Length 3014;
  Best Local Similarity   88.5%;  Pred. No. 0;
  Matches 2474;  Conservative    0;  Mismatches  312;  Indels   11;  Gaps    4;
```

```
Qy         87 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 146

Db          1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGTGCTCCTGGGCATAGTCCTCGCCGCCTGGCTT 60

Qy        147 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 206

Db         61 CGTGGTTCGGGTGCCCAGCAGAGTGCCACGGTGGCCAATCCAGTGCCCGGTGCCAACCCC 120

Qy        207 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 266

Db        121 GACCTGCTGCCCCACTTCCTGGTAGAGCCTGAGGACGTGTACATTGTCAAGAACAAGCCG 180

Qy        267 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 326

Db        181 GTGTTGTTGGTGTGCAAGGCTGTGCCTGCCACCCAGATCTTCTTCAAGTGCAATGGGGAA 240

Qy        327 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCG 386

Db        241 TGGGTCCGCCAGGTCGATCACGTAATTGAACGCAGCACCGACAGCAGCAGCGGATTGCCA 300

Qy        387 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 446

Db        301 ACCATGGAGGTCCGTATCAACGTATCGAGGCAGCAGGTAGAGAAAGTGTTTGGGCTGGAG 360

Qy        447 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 506

Db        361 GAATACTGGTGCCAGTGTGTGGCATGGAGCTCCTCGGGTACCACCAAAAGTCAGAAGGCC 420

Qy        507 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 566

Db        421 TACATCCGGATTGCCTATTTGCGCAAGAACTTTGAGCAGGAGCCACTGGCCAAGGAAGTG 480

Qy        567 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 626

Db        481 TCACTGGAGCAAGGCATTGTACTACCTTGTCGCCCCCCAGAAGGAATCCCCCCAGCTGAG 540

Qy        627 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 686

Db        541 GTGGAGTGGCTTCGAAATGAGGACCTCGTGGACCCCTCCCTCGATCCCAATGTGTACATC 600

Qy        687 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 746

Db        601 ACGCGGGAGCACAGCCTAGTCGTGCGTCAGGCCCGCCTGGCCGACACGGCCAACTACACC 660

Qy        747 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 806

Db        661 TGTGTGGCCAAGAACATCGTAGCCCGTCGCCGAAGCACCTCTGCAGCGGTCATTGTTTAT 720

Qy        807 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 866

Db        721 GTGAACGGTGGGTGGTCGACGTGGACTGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGT 780

Qy        867 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 926

Db        781 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCACCTCTCAACGGGGGCGCCTTC 840
```

```
Qy        927 TGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGC 986
Db        841 TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACTCTG---TGCCCAGTGGATGGG 897

Qy        987 AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC 1046
Db        898 AGCTGGAGTTCGTGGAGTAAGTGGTCAGCCTGTGGGCTTGACTGCACCCACTGGCGGAGC 957

Qy       1047 CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG 1106
Db        958 CGCGAGTGCTCTGACCCAGCACCCCGCAATGGAGGTGAGGAGTGTCGGGGTGCTGACCTG 1017

Qy       1107 GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG 1166
Db       1018 GACACCCGCAACTGTACCAGTGACCTCTGCCTGCACACCGCTTCTTGCCCCGAGGACGTG 1077

Qy       1167 GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC 1226
Db       1078 GCTCTCTACATCGGCCTTGTCGCTGTGGCTGTGTGCCTCTTCTTGCTGTTGCTGGCCCTT 1137

Qy       1227 ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1286
Db       1138 GGACTCATTTACTGTCGCAAGAAGGAAGGGCTGGACTCCGATGTGGCCGACTCGTCCATC 1197

Qy       1287 CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1346
Db       1198 CTCACCTCGGGCTTCCAGCCTGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCACCTG 1257

Qy       1347 CTCACCATCCAGCCGGACCTCAGCACCACC---ACCACCTACCAGGGCAGTCTCTGTCCC 1403
Db       1258 CTCACCATCCAGCCAGACCTCAGCACCACCACTACCACCTACCAGGGCAGTCTATGTTCG 1317

Qy       1404 CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG 1463
Db       1318 AGGCAGGATGGACCCAGCCCCAAGTTCCAGCTCTCTAATGGTCACCTGCTCAGCCCACTG 1377

Qy       1464 GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC 1523
Db       1378 GGGAGTGGCCGCCATACGTTGCACCACAGCTCACCCACCTCTGAGGCTGAGGACTTCGTC 1437

Qy       1524 TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC 1583
Db       1438 TCCCGCCTCTCCACCCAAAACTACTTTCGTTCCCTGCCCCGCGGCACCAGCAACATGGCC 1497

Qy       1584 TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC 1643
Db       1498 TACGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACGGGGATCAGCCTC 1557

Qy       1644 CTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC 1703
Db       1558 CTCATACCCCCGGATGCCATCCCCCGAGGAAAGATCTACGAGATCTACCTCACACTGCAC 1617

Qy       1704 AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT 1763
Db       1618 AAGCCAGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCAGTCGTT 1677

Qy       1764 AGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT 1823
Db       1678 AGCTGTGGGCCCCCAGGAGTCCTGCTCACCCGGCCAGTCATCCTTGCAATGGACCACTGT 1737

Qy       1824 GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC 1883
Db       1738 GGAGAGCCCAGCCCTGACAGCTGGAGTCTGCGCCTCAAAAAGCAGTCCTGCGAGGGCAGT 1797

Qy       1884 TGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG 1943
Db       1798 TGGG---AGGATGTGCTGCACCTTGGTGAGGAGTCACCTTCCCACCTCTACTACTGCCAG 1854

Qy       1944 CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA 2003
Db       1855 CTGGAGGCCGGGGCCTGCTATGTCTTCACGGAGCAGCTGGGCCGCTTTGCCCTGGTAGGA 1914

Qy       2004 GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC 2063
Db       1915 GAGGCCCTCAGCGTGGCTGCCACCAAGCGCCTCAGGCTCCTTCTGTTTGCTCCCGTGGCC 1974

Qy       2064 TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC 2123
Db       1975 TGTACGTCCCTTGAGTACAACATCCGAGTGTACTGCCTACACGACACCCACGACGCTCTC 2034

Qy       2124 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTC 2183
Db       2035 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTAGGTGGACAGCTGATCCAGGAGCCTCGCGTC 2094

Qy       2184 CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCC 2243
Db       2095 CTGCACTTCAAAGACAGTTACCACAACCTACGTCTCTCCATCCACGACGTGCCCAGCTCC 2154

Qy       2244 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2303
Db       2155 CTGTGGAAGAGCAAGCTACTTGTCAGCTACCAGGAGATCCCTTTTTACCACATCTGGAAC 2214

Qy       2304 GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2363
Db       2215 GGCACCCAGCAGTATCTGCACTGCACCTTCACCCTGGAGCGCATCAACGCCAGCACCAGC 2274

Qy       2364 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGCAGAGCTTCAGCATC 2423
Db       2275 GACCTGGCCTGCAAGGTGTGGGTGTGGCAGGTGGAGGGAGATGGGCAGAGCTTCAACATC 2334

Qy       2424 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2483
Db       2335 AACTTCAACATCACTAAGGACACAAGGTTTGCTGAATTGTTGGCTCTGGAGAGTGAAGGG 2394

Qy       2484 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2543
Db       2395 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAAAAG 2454

Qy       2544 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2603
Db       2455 ATCATCGCCAGTCTGGACCCACCCTGCAGCCGGGGCGCCGACTGGAGAACTCTAGCCCAG 2514

Qy       2604 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG 2663
Db       2515 AAACTTCACCTGGACAGCCATCTTAGCTTCTTTGCCTCCAAGCCCAGCCCTACAGCCATG 2574

Qy       2664 ATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2723
Db       2575 ATCCTCAACCTATGGAGGCACGGCACTTCCCCAACGGCAACCTCGGCCAGCTGGCAGCA 2634

Qy       2724 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2783
Db       2635 GCTGTGGCCGGACTGGGCCAACCAGATGCTGGCCTCTTCACGGTGTCGGAGGCCGAGTGT 2694

Qy       2784 TGAGGCCGGCCAGGCCCG--ACACCTACACTCTCACCAGCTTTGGCACCCACCAAGGACA 2841
Db       2695 TGAGACCAGCCAGGCCGGTAATGCCTACATTCTCACCAGCTTTGACACCTGCCAGGGACA 2754

Qy       2842 GGCAGAAGCCGGACAGGGGCCCTTCCCCACACCGGGG 2878
Db       2755 GGCAAAACCAGACAGGGGCCCTTCCCCCACACCCGGG 2791
```

RESULT 3
US-08-808-982-2
; Sequence 2, Application US/08808982
; Patent No. 5939271
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; APPLICANT: Leonardo, E. David
; APPLICANT: Hink, Lindsay
; APPLICANT: Masu, Masayuki
; APPLICANT: Kazuko, Keino-Masu
; TITLE OF INVENTION: Netrin Receptors
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:

```
;       ADDRESSEE:  SCIENCE & TECHNOLOGY LAW GROUP
;       STREET:  268 BUSH STREET, SUITE 3200
;       CITY:  SAN FRANCISCO
;       STATE:  CALIFORNIA
;       COUNTRY:  USA
;       ZIP:  94104
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/808,982
;       FILING DATE:
;       CLASSIFICATION:  530
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  OSMAN, RICHARD A
;       REGISTRATION NUMBER:  36,627
;       REFERENCE/DOCKET NUMBER:  UC96-217
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (415) 343-4341
;       TELEFAX:  (415) 343-4342
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1787 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
US-08-808-982-2

  Query Match             54.5%;  Score 1570.4;  DB 2;  Length 1787;
  Best Local Similarity   96.9%;  Pred. No. 0;
  Matches 1721;  Conservative    0;  Mismatches   36;  Indels   19;  Gaps   11;

Qy       1114 GCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCT 1173
Db          1 GCAACTGTACCAGTGACCTCTG-GTACACACTGCTTCTGGCCCTGAGGACGTGGCCCTCT 59

Qy       1174 ATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCATCCTCG 1233
Db         60 ATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCATCCTCG 119

Qy       1234 TTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCT 1293
Db        120 TTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCT 179

Qy       1294 CAGGCTTCCAGCCCGTCAGCATC-AAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACC 1352
Db        180 CAGGCTTCCAGCCCGTCAGCATCTAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACC 239

Qy       1353 ATCCAGCCGGACCTCAG---CACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAG 1409
Db        240 ATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAG 299

Qy       1410 GATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGC 1469
Db        300 GATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGC 359

Qy       1470 GGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGC 1529
Db        360 GGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGC 419

Qy       1530 CTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGG 1589
Db        420 CTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGG 479

Qy       1590 ACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATC 1649
Db        480 ACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGAATCAGCCTCCTCATC 539

Qy       1650 CCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCG 1709
Db        540 CCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCG 599

Qy       1710 GAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGT 1769
Db        600 GAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGT 659

Qy       1770 GGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAG 1829
Db        660 GGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAG 719

Qy       1830 CCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAG 1889
Db        720 CCCAGCCCTGACAGCTGGAGCCTGGCCCTCAAAAAGCAGTCGTGCGAGGG-AGCTGGGAG 778

Qy       1890 CAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG 1949
Db        779 GATG----TCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG 834

Qy       1950 GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC 2009
Db        835 GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC 894

Qy       2010 CTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC 2069
Db        895 CTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC 954

Qy       2070 TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG 2129
Db        955 TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG 1014

Qy       2130 GTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC 2189
Db       1015 GTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC 1074

Qy       2190 TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG 2249
Db       1075 TT-AAGGACAGTTACCACAACCT--GCCCTATCATCCACGATGTGCCCAGCTCCCTGTGG 1131

Qy       2250 AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG 2309
Db       1132 AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG 1191

Qy       2310 CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG 2369
Db       1192 CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG 1251

Qy       2370 GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC 2429
Db       1252 GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC 1311

Qy       2430 AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC 2489
Db       1312 AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC 1371

Qy       2490 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT 2549
Db       1372 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT 1431

Qy       2550 TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC 2609
Db       1432 TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC 1491

Qy       2610 CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC 2669
Db       1492 CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC 1551

Qy       2670 AACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG 2729
Db       1552 AACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG 1611

Qy       2730 GCTGGACTGGGCCAGCCAGACGCTGGCCTC-TTCACAGTG-TCGGAGGCTGAGTGCTGAG 2787
Db       1612 GCTGGACTGGCCAGCAGGACGGTGGCTTCTTTCACAGTGTTCGGAGGCTGAGTGCTGAG 1671
```

```
Qy       2788 GCCGGCCAGGCCCGACACCTACACTCTCACCAGCTTTGGCA--CCCACCAAGGACAGGCA 2845
              ||||||||||  |||    ||||| | | ||||| |||||  |   ||   | | ||||||||
Db       1672 GCCGGCCAGG--CGAACACTACAATTTTACCAGTTTGGGAACCCACCAAGGACAGGCA 1729

Qy       2846 GAAGCCGGACAGGGGCCCTTCCCCACACCGGGGAGA 2881
              |||||||||| ||||   ||||| | |||||||||
Db       1730 GAAGCCGGACAAGGGCTTTTCCCAAAACCGGGGAGA 1765


RESULT 4
US-09-306-902A-2
; Sequence 2, Application US/09306902A
; Patent No. 6277585
;  GENERAL INFORMATION:
;       APPLICANT: Tessier-Lavigne, Marc
;                  Leonardo, E. David
;                  Hink, Lindsay
;                  Masu, Masayuki
;                  Kazuko, Keino-Masu
;       TITLE OF INVENTION: Netrin Receptors
;       NUMBER OF SEQUENCES: 9
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
;            STREET: 268 BUSH STREET, SUITE 3200
;            CITY: SAN FRANCISCO
;            STATE: CALIFORNIA
;            COUNTRY: USA
;            ZIP: 94104
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/306,902A
;            FILING DATE: 07-May-1999
;            CLASSIFICATION: &lt;Unknown&gt;
;       ATTORNEY/AGENT INFORMATION:
;            NAME: OSMAN, RICHARD A
;            REGISTRATION NUMBER: 36,627
;            REFERENCE/DOCKET NUMBER: UC96-217
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (415) 343-4341
;            TELEFAX: (415) 343-4342
;  INFORMATION FOR SEQ ID NO: 2:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 1787 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: double
;            TOPOLOGY: linear
;       MOLECULE TYPE: cDNA
;       SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-306-902A-2

  Query Match             54.5%;  Score 1570.4;  DB 3;  Length 1787;
  Best Local Similarity   96.9%;  Pred. No. 0;
  Matches 1721;  Conservative    0;  Mismatches   36;  Indels   19;  Gaps   11;

Qy       1114 GCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCT 1173
              |||||||||||||||||||||| ||||||| ||||||||||||||||||||||||||||||
Db          1 GCAACTGTACCAGTGACCTCTG-GTACACACTGCTTCTGGCCCTGAGGACGTGGCCCTCT 59

Qy       1174 ATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCATCCTCG 1233
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         60 ATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCATCCTCG 119

Qy       1234 TTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCT 1293
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        120 TTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCT 179

Qy       1294 CAGGCTTCCAGCCCGTCAGCATC-AAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACC 1352
              ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        180 CAGGCTTCCAGCCCGTCAGCATCTAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACC 239

Qy       1353 ATCCAGCCGGACCTCAG---CACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAG 1409
              |||||||||||||||||   ||||||||||||||||||||||||||||||||||||||||
Db        240 ATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAG 299

Qy       1410 GATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGC 1469
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        300 GATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGC 359

Qy       1470 GGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGC 1529
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        360 GGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGC 419

Qy       1530 CTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGG 1589
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        420 CTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGG 479

Qy       1590 ACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATC 1649
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        480 ACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGAATCAGCCTCCTCATC 539

Qy       1650 CCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCG 1709
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 CCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCG 599

Qy       1710 GAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGT 1769
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 GAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGT 659

Qy       1770 GGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAG 1829
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 GGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAG 719

Qy       1830 CCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAG 1889
              ||||||||||||||||||||||| ||||||||||||||||||||||||| ||||||||
Db        720 CCCAGCCCTGACAGCTGGAGCCTGGCCCTCAAAAAGCAGTCGTGCGAGGG-AGCTGGGAG 778

Qy       1890 CAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG 1949
              | |    |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        779 GATG----TCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG 834

Qy       1950 GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC 2009
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        835 GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC 894

Qy       2010 CTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC 2069
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        895 CTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC 954

Qy       2070 TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG 2129
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        955 TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG 1014

Qy       2130 GTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC 2189
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1015 GTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC 1074

Qy       2190 TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG 2249
              || |||||||||||||||||  |||    ||||||||||||||||||||||||||||||
Db       1075 TT-AAGGACAGTTACCACAACCT--GCCCTATCATCCACGATGTGCCCAGCTCCCTGTGG 1131

Qy       2250 AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG 2309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1132 AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG 1191

Qy       2310 CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG 2369
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1192 CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG 1251

Qy       2370 GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC 2429
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1252 GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC 1311
```

```
Qy   2430 AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC 2489
Db   1312 AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC 1371

Qy   2490 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT 2549
Db   1372 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT 1431

Qy   2550 TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC 2609
Db   1432 TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC 1491

Qy   2610 CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC 2669
Db   1492 CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC 1551

Qy   2670 AACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG 2729
Db   1552 AACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG 1611

Qy   2730 GCTGGACTGGGCCAGCCAGACGCTGGCCTC-TTCACAGTG-TCGGAGGCTGAGTGCTGAG 2787
Db   1612 GCTGGGACTGGCCAGCAGGACGGTGGCTTCTTTCACAGTGTTCGGAGGCTGAGTGCTGAG 1671

Qy   2788 GCCGGCCAGGCCCGACACCTACACTCTCACCAGCTTTGGCA--CCCACCAAGGACAGGCA 2845
Db   1672 GCCGGCCAGG--CGAACACTACAATTTTACCAGTTTTGGGAACCCACCAAGGGACAGGCA 1729

Qy   2846 GAAGCCGGACAGGGGCCCTTCCCCACACCGGGGAGA 2881
Db   1730 GAAGCCGGACAAGGGCTTTTCCCAAAACCGGGGAGA 1765
```

RESULT 5
US-09-833-381-1806
; Sequence 1806, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1806
; LENGTH: 1282
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1282)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1806

```
  Query Match             39.8%;  Score 1147.8;  DB 4;  Length 1282;
  Best Local Similarity   97.7%;  Pred. No. 3.3e-240;
  Matches 1259;  Conservative    0;  Mismatches   18;  Indels   11;  Gaps    9;

Qy   1510 CCGAGGAGTTCGTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCA 1569
Db      1 CCGAGGAGTTCGTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCA 60

Qy   1570 CCAGCAACATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATA 1629
Db     61 CCAGCAACATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATA 120

Qy   1630 CAGGTATCAGCCTCCTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCT 1689
Db    121 CAGGAATCAGCCTCCTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCT 180

Qy   1690 ACCTCACGCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGC 1749
Db    181 ACCTCACGCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGC 240

Qy   1750 TGAGTCCCATCGTTAGCTGTGGACCCCCT-GGCGTCCTGCTCACCCGGCCAGTCATCCT- 1807
Db    241 TGAGTCCCATCGTTAGCTGTGGACCCCCTGGGCGTCCTGCTCACCCGGCCAGTCATCCTG 300

Qy   1808 GGCTATGGACCACTGTGGGGAGCCCAGCCCTGACAGCT-GGAGCCTGCGCCTCAAAAAGC 1866
Db    301 GGGTATGGACCACTGTGGGGAGCCCAGCCCTGACAGCTGGGAGCCTGCGCCTCAAAAAGC 360

Qy   1867 AGTCGTGCGAGGGCAGCTGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCC 1926
Db    361 AGTCGTGCGAGGGCAGCTGGGAGGATGTGCT--TGCACCTGGGCGAGGAGGCGCCCTCCC 418

Qy   1927 ACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCC 1986
Db    419 ACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGAGCC 478

Qy   1987 GCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTC 2046
Db    479 GCTATGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTC 538

Qy   2047 TGTTTGCGCCGGTGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATG 2106
Db    539 TGTTTGCGCCGGTGGCCTGCACCTCCCTCGAGTACAACATACTGGTCTACTGCCTGCATG 598

Qy   2107 ACACCCACGATGCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGA 2166
Db    599 ACACTCACGATGCACTCAACGTAGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGA 658

Qy   2167 TCCAGGAGCCACGGGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCC 2226
Db    659 TCCAGGAGCCACGGGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCC 718

Qy   2227 ACGATGTGCCCAGCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCT 2286
Db    719 ACGATGTGCCCAGCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCT 778

Qy   2287 TTTATCACATCTGGAATGGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTG 2346
Db    779 TTTATCACATCTGGAATGGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTG 838

Qy   2347 TCAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACG 2406
Db    839 TCAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACG 898

Qy   2407 GGCAGAGCTTCAGCATCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGG 2466
Db    899 GGCAGAGCTTCAGCATCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGG 958

Qy   2467 CTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCT 2526
Db    959 CTCTGGAGAGTGAAGCGNGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCT 1018

Qy   2527 TCCTCATTCGGCAGAAGATAA-TTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGAC 2585
Db   1019 TCCTCATTCGGCAGAAGATAATTTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGAC 1078

Qy   2586 TGGCGG-ACTCTGGCCCAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAA 2644
Db   1079 TGGCGGAACTCTGGCCCAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAA 1138

Qy   2645 GCCCAGCCCCACAGCCATGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAA 2704
Db   1139 GCCCAGCCCCACAGCCATGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAA 1198

Qy   2705 CCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCAC 2764
Db   1199 CCTCAGCCAGCTGG-TGCAGCAGTGGCTGGACTGGGC--AGCAGACGCTGG-CTCTTCAC 1254
```

```
Qy       2765 AGTGTCGGAGGCTGAGTGCTGAGGCCGG 2792
Db       1255 AGTGTCGGAGGCTGAGTGCTGAGGCCGG 1282

RESULT 6
US-08-808-982-3
; Sequence 3, Application US/08808982
; Patent No. 5939271
;  GENERAL INFORMATION:
;    APPLICANT:  Tessier-Lavigne, Marc
;    APPLICANT:  Leonardo, E. David
;    APPLICANT:  Hink, Lindsay
;    APPLICANT:  Masu, Masayuki
;    APPLICANT:  Kazuko, Keino-Masu
;    TITLE OF INVENTION:  Netrin Receptors
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SCIENCE & TECHNOLOGY LAW GROUP
;      STREET:  268 BUSH STREET, SUITE 3200
;      CITY:  SAN FRANCISCO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/808,982
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  OSMAN, RICHARD A
;      REGISTRATION NUMBER:  36,627
;      REFERENCE/DOCKET NUMBER:  UC96-217
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 343-4341
;      TELEFAX:  (415) 343-4342
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2831 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-808-982-3

  Query Match             27.2%;  Score 783.4;  DB 2;  Length 2831;
  Best Local Similarity   59.4%;  Pred. No. 5.7e-161;
  Matches 1625;  Conservative    0;  Mismatches  971;  Indels  139;  Gaps   12;

Qy        184 ACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATG 243
Db        104 ACTCCTTCCCATCAGCACCCGCGGAGCAGCTGCCTCACTTCCTGCTGGAACCAGAGGATG 163

Qy        244 TGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGA 303
Db        164 CCTACATCGTAAAGAACAAGCCAGTGGAATTGCACTGCCGAGCCTTCCCTGCCACACAGA 223

Qy        304 TCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCA 363
Db        224 TCTACTTCAAGTGTAATGGCGAGTGGGTTAGCCAGAAAGGCCACGTCACGCAGGAGAGCC 283

Qy        364 CAGACGGGAGCAGTGGTGAGCCGACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGG 423
Db        284 TGGATGAGGCCACAGGCTTGCGAATACGAGAGGTGCAGATAGAGGTGTCGCGGCAGCAGG 343

Qy        424 TCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGG 483
Db        344 TGGAGGAACTTTTTGGGCTCGAGGACTACTGGTGTCAGTGCGTGGCCTGGAGCTCTTCGG 403

Qy        484 GCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGC 543
Db        404 GAACCACCAAGAGTCGCCGAGCCTACATCCGCATTGCCTACTTGCGCAAGAACTTTGACC 463

Qy        544 AGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCAC 603
Db        464 AGGAGCCTCTGGCGAAGGAGGTACCCTTGGATCATGAGGTCCTTCTGCAGTGCCGCCCAC 523

Qy        604 CGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGT 663
Db        524 CAGAGGGAGTGCCTGTGGCTGAGGTGGAATGGCTCAAGAATGAAGATGTCATCGATCCCG 583

Qy        664 CCCTGGACCCCAATGTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCC 723
Db        584 CTCAGGACACTAACTTCCTGCTCACCATTGACCACAACCTCATCATCCGCCAGGCGCGCC 643

Qy        724 TTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCG 783
Db        644 TCTCAGACACAGCCAACTACACCTGTGTGGCAAAGAATATTGTGGCCAAGCGCCGGAGCA 703

Qy        784 CCTCCGCTGCTGTCATCGTCTACGTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCG 843
Db        704 CGACGGCCACAGTCATCGTCTATGTGAACGGAGGTTGGTCCAGCTGGGCAGAATGGTCAC 763

Qy        844 TCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGG 903
Db        764 CCTGCTCTAACCGCTGCGGCCGAGGTTGGCAGAAACGTACTAGGACCTGCACCAACCCAG 823

Qy        904 CGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCCT 963
Db        824 CCCCACTCAATGGAGGTGCCTTCTGCGAGGGACAGGCTTGCCAGAA---GACGGCTTGCA 880

Qy        964 CTCTGCTTGTCTCTGTGGACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGC 1023
Db        881 CCACCGTGTGCCCAGTGGATGGAGCGTGGACTGAGTGGAGCAAGTGGTCCGCCTGCAGCA 940

Qy       1024 TGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGG 1083
Db        941 CAGAGTGTGCGCACTGGCGCAGCCGCGAGTGCATGGCACCGCCGCCCCAGAACGGAGGCC 1000

Qy       1084 AGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGTACCAGTGACCTCTGTGTACACA 1143
Db       1001 GTGACTGCAGCGGGACGCTACTTGACTCCAAGAACTGCACCGATGGGCTGTGCGTGCTGA 1060

Qy       1144 GTGCTTCTGGCC------------------------------CTGAGGACGTGGCCC 1170
Db       1061 ATCAGAGAACTCTAAACGACCCTAAAAGCCGCCCCCTGGAGCCGTCGGGAGACGTGGCGC 1120

Qy       1171 TCTATGTGGGCCTC---ATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCA 1227
Db       1121 TGTATGCGGGCCTCGTGGTGGCCGTCTTTGTGGTTCTGGCAGTTCTCATGGCTGTAGGAG 1180

Qy       1228 TCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT- 1286
Db       1181 TGATCGTGTACCGGAGAAACTGCCGGGACTTCGACACGGACATCACTGACTCCTCTGCTG 1240

Qy       1287 --CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATC 1344
Db       1241 CCCTCACTGGTGGTTTCCACCCCGTCAACTTCAAGACTGCAAGGCCCAGCAACCCACAGC 1300

Qy       1345 TGCT------CACCATCCAGCCGGACCTCAGCACCA---CCACCACCTACCAGGGCAGTC 1395
Db       1301 TCCTGCACCCATCCGCCCCTCCGGACCTAACGGCCAGTGCTGGCATCTACCGCGGACCTG 1360

Qy       1396 TCTGTCCCCGGCAGGA------------------------------------------- 1411
Db       1361 TGTATGCCCTGCAGGACTCTGCCGACAAGATCCCTATGACTAATTCACCCCTTCTGGATC 1420

Qy       1412 -TGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCC---------CC 1461
Db       1421 CCTTGCCCAGCCTCAAGATCAAGGTCTATGACTCCAGCACCATCGGCTCTGGGGCTGGCC 1480

Qy       1462 TGGGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCG 1521
```

```
Db   1481 TGGCTGATGGAGCCGACCTGCTGGGTGTCTTACCACCCGGTACATACCCAGGCGATTTCT 1540
Qy   1522 TCTCCCGCCTCTCCACCCAGAAC--------------------------TACTTCCGCT 1554
Db   1541 CCCGGGACACCCACTTCCTGCACCTGCGCAGCGCCAGCCTTGGTTCCCAGCACCTCCTGG 1600
Qy   1555 CCCTGCCCCGAGGCACCAGCAACATGACCTATGGACCTTCAACTTCCTCGGGGGCCGGC 1614
Db   1601 GCCTCCCTCGAGACCCCAGCAGCAGTGTCAGTGGCACCTTTGGTTGCCTGGGTGGGAGGC 1660
Qy   1615 TGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCCAGATGCCATACCCCGAGGGA 1674
Db   1661 TGACCATTCCCGGCACAGGGGTCAGCCTGTTGGTACCAAATGGAGCCATTCCCCAGGGCA 1720
Qy   1675 AGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTG 1734
Db   1721 AGTTCTATGACTTGTATCTACGTATCAACAAGACTGAAAGCACCCTCCCACTTTCGGAAG 1780
Qy   1735 GCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCTGGCGTCCTGCTCACCC 1794
Db   1781 GTTCCCAGACAGTATTGAGCCCCTCGGTGACCTGCGGGCCCACGGGCCTCCTCCTGTGCC 1840
Qy   1795 GGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGC 1854
Db   1841 GCCCTGTTGTCCTCACTGTGCCCCACTGTGCTGAAGTCATTGCCGGAGACTGGATCTTCC 1900
Qy   1855 GCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGCAGGATGTGCTGCACCTGGGCGAGG 1914
Db   1901 AGCTCAAGACCCAGGCCCATCAGGGCCACTGGG---AGGAGGTGGTGACTTTGGATGAGG 1957
Qy   1915 AGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCG 1974
Db   1958 AGACTCTGAACACCCCCTGCTACTGCCAGCTAGAGGCTAAATCCTGCCACATCCTGTTGG 2017
Qy   1975 AGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCC 2034
Db   2018 ACCAGCTGGGTACCTACGTGTTCACGGGCGAGTCCTACTCCCGCTCCGCAGTCAAGCGGC 2077
Qy   2035 TCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCT 2094
Db   2078 TCCAGCTAGCCATCTTCGCCCCAGCCCTCTGCACCTCCCTGGAGTATAGTCTCAGGGTCT 2137
Qy   2095 ACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGG 2154
Db   2138 ACTGTCTGGAGGACACTCCTGCAGCACTGAAGGAGGTCCTAGAGCTGGAGAGGACTCTGG 2197
Qy   2155 GGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAGTTACCACAACCTGC 2214
Db   2198 GTGGCTACTTGGTGGAGGAGCCCAAGACTTTGCTCTTTAAGGACAGTTACCACAACCTAC 2257
Qy   2215 GCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACC 2274
Db   2258 G-CTCTCCCTCCATGACATCCCCCATGCCCACTGGAGGAGCAAACTACTGGCCAAGTACC 2316
Qy   2275 AGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTTGCACTGCACCTTCA 2334
Db   2317 AGGAGATTCCCTTCTACCATGTGTGGAACGGCAGCCAGAAAGCCCTGCACTGCACTTTCA 2376
Qy   2335 CCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGG 2394
Db   2377 CCCTGGAGAGACATAGCCTAGCCTCCACTGAGTTCACCTGTAAGGTCTGCGTGCGGCAGG 2436
Qy   2395 TGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCAC---CAAGGACACAAGGT 2451
Db   2437 TAGAAGGGGAAGGCCAGATTTTCCAGCTGCACACCACGCTGGCTGAGACGCCTGCTGGCT 2496
Qy   2452 TTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTG 2511
Db   2497 CCCTGGATGCACTCTGCTCTGCCCCTGGCAATGCTGCCACCACACAGCTGGGACCCTATG 2556
Qy   2512 CCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTGGACCCACCCTGTA 2571
Db   2557 CCTTCAAGATACCACTGTCCATCCGCCAGAAGATCTGCAACAGCCTGGACGCCCCCAACT 2616
Qy   2572 GGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGACAGCCATCTCAGCT 2631
Db   2617 CACGGGGCAATGACTGGCGGCTGTTGGCACAGAAGCTCTCCATGGACCGGTACCTGAACT 2676
Qy   2632 TCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGGGAGGCGCGGCACT 2691
Db   2677 ACTTCGCCACCAAAGCTAGTCCCACAGGCGTGATCTTAGACCTCTGGGAAGCTCGGCAGC 2736
Qy   2692 TCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGGCCAGCCAGACG 2751
Db   2737 AGGATGATGGGGACCTCAACAGCCTGGCCAGTGCCTTGGAGGAGATGGGCAAGAGTGAGA 2796
Qy   2752 CTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2786
Db   2797 TGCTGGTAGCCATGACCACTGATGGCGATTGCTGA 2831
```

RESULT 7
US-09-306-902A-3
; Sequence 3, Application US/09306902A
; Patent No. 6277585
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; Leonardo, E. David
; Hink, Lindsay
; Masu, Masayuki
; Kazuko, Keino-Masu
; TITLE OF INVENTION: Netrin Receptors
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/306,902A
; FILING DATE: 07-May-1999
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2831 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-306-902A-3

Query Match 27.2%; Score 783.4; DB 3; Length 2831;
Best Local Similarity 59.4%; Pred. No. 5.7e-161;
Matches 1625; Conservative 0; Mismatches 971; Indels 139; Gaps 12;

```
Qy    184 ACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATG 243
Db    104 ACTCCTTCCCATCAGCACCCGCGGAGCAGCTGCCTCACTTCCTGCTGGAACCAGAGGATG 163
Qy    244 TGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGA 303
```

```
Db    164 CCTACATCGTAAAGAACAAGCCAGTGGAATTGCACTGCCGAGCCTTCCCTGCCACACAGA 223

Qy    304 TCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCA 363
Db    224 TCTACTTCAAGTGTAATGGCGAGTGGGTTAGCCAGAAAGGCCACGTCACGCAGGAGAGCC 283

Qy    364 CAGACGGGAGCAGTGGTGAGCCGACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGG 423
Db    284 TGGATGAGGCCACAGGCTTGCGAATACGAGAGGTGCAGATAGAGGTGTCGCGGCAGCAGG 343

Qy    424 TCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGG 483
Db    344 TGGAGGAACTTTTTGGGCTCGAGGACTACTGGTGTCAGTGCGTGGCCTGGAGCTCTTCGG 403

Qy    484 GCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGC 543
Db    404 GAACCACCAAGAGTCGCCGAGCCTACATCCGCATTGCCTACTTGCGCAAGAACTTTGACC 463

Qy    544 AGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCAC 603
Db    464 AGGAGCCTCTGGCGAAGGAGGTACCCTTGGATCATGAGGTCCTTCTGCAGTGCCGCCCAC 523

Qy    604 CGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGT 663
Db    524 CAGAGGGAGTGCCTGTGGCTGAGGTGGAATGGCTCAAGAATGAAGATGTCATCGATCCCG 583

Qy    664 CCCTGGACCCCAATGTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCC 723
Db    584 CTCAGGACACTAACTTCCTGCTCACCATTGACCACAACCTCATCATCCGCCAGGCGCGCC 643

Qy    724 TTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCG 783
Db    644 TCTCAGACACAGCCAACTACACCTGTGTGGCAAAGAATATTGTGGCCAAGCGCCGGAGCA 703

Qy    784 CCTCCGCTGCTGTCATCGTCTACGTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCG 843
Db    704 CGACGGCCACAGTCATCGTCTATGTGAACGGAGGTTGGTCCAGCTGGGCAGAATGGTCAC 763

Qy    844 TCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGG 903
Db    764 CCTGCTCTAACCGCTGCGGCCGAGGTTGGCAGAAACGTACTAGGACCTGCACCAACCCAG 823

Qy    904 CGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCCT 963
Db    824 CCCCACTCAATGGAGGTGCCTTCTGCGAGGGACAGGCTTGCCAGAA---GACGGCTTGCA 880

Qy    964 CTCTGCTTGTCTCTGTGGACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGC 1023
Db    881 CCACCGTGTGCCCAGTGGATGGAGCGTGGACTGAGTGGAGCAAGTGGTCCGCCTGCAGCA 940

Qy   1024 TGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGG 1083
Db    941 CAGAGTGTGCGCACTGGCGCAGCCGCGAGTGCATGGCACCGCCGCCCCAGAACGGAGGCC 1000

Qy   1084 AGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGTACCAGTGACCTCTGTGTACACA 1143
Db   1001 GTGACTGCAGCGGGACGCTACTTGACTCCAAGAACTGCACCGATGGGCTGTGCGTGCTGA 1060

Qy   1144 GTGCTTCTGGCC-----------------------------------CTGAGGACGTGGCCC 1170
Db   1061 ATCAGAGAACTCTAAACGACCCTAAAAGCCGCCCCCTGGAGCCGTCGGGAGACGTGGCGC 1120

Qy   1171 TCTATGTGGGCCTC---ATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCA 1227
Db   1121 TGTATGCGGGCCTCGTGGTGGCCGTCTTTGTGGTTCTGGCAGTTCTCATGGCTGTAGGAG 1180

Qy   1228 TCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT- 1286
Db   1181 TGATCGTGTACCGGAGAAACTGCCGGGACTTCGACACGGACATCACTGACTCCTCTGCTG 1240

Qy   1287 --CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATC 1344
Db   1241 CCCTCACTGGTGGTTTCCACCCCGTCAACTTCAAGACTGCAAGGCCCAGCAACCCACAGC 1300

Qy   1345 TGCT------CACCATCCAGCCGGACCTCAGCACCA---CCACCACCTACCAGGGCAGTC 1395
Db   1301 TCCTGCACCCATCCGCCCCTCCGGACCTAACGGCCAGTGCTGGCATCTACCGCGGACCTG 1360

Qy   1396 TCTGTCCCCGGCAGGA-------------------------------------------- 1411
Db   1361 TGTATGCCCTGCAGGACTCTGCCGACAAGATCCCTATGACTAATTCACCCCTTCTGGATC 1420

Qy   1412 -TGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCC---------CC 1461
Db   1421 CCTTGCCCAGCCTCAAGATCAAGGTCTATGACTCCAGCACCATCGGCTCTGGGGCTGGCC 1480

Qy   1462 TGGGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCG 1521
Db   1481 TGGCTGATGGAGCCGACCTGCTGGGTGTCTTACCACCCGGTACATACCCAGGCGATTTCT 1540

Qy   1522 TCTCCCGCCTCTCCACCCAGAAC-------------------------TACTTCCGCT 1554
Db   1541 CCCGGGACACCCACTTCCTGCACCTGCGCAGCGCCAGCCTTGGTTCCCAGCACCTCCTGG 1600

Qy   1555 CCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGC 1614
Db   1601 GCCTCCCTCGAGACCCCAGCAGCAGTGTCAGTGGCACCTTTGGTTGCCTGGGTGGGAGGC 1660

Qy   1615 TGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCCAGATGCCATACCCCGAGGGA 1674
Db   1661 TGACCATTCCCGGCACAGGGGTCAGCCTGTTGGTACCAAATGGAGCCATTCCCCAGGGCA 1720

Qy   1675 AGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTG 1734
Db   1721 AGTTCTATGACTTGTATCTACGTATCAACAAGACTGAAAGCACCCTCCCACTTTCGGAAG 1780

Qy   1735 GCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCTGGCGTCCTGCTCACCC 1794
Db   1781 GTTCCCAGACAGTATTGAGCCCCTCGGTGACCTGCGGGCCCACGGGCCTCCTCCTGTGCC 1840

Qy   1795 GGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGC 1854
Db   1841 GCCCTGTTGTCCTCACTGTGCCCCACTGTGCTGAAGTCATTGCCGGAGACTGGATCTTCC 1900

Qy   1855 GCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGCAGGATGTGCTGCACCTGGGCGAGG 1914
Db   1901 AGCTCAAGACCCAGGCCCATCAGGGCCACTGGG---AGGAGGTGGTGACTTTGGATGAGG 1957

Qy   1915 AGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCG 1974
Db   1958 AGACTCTGAACACCCCCTGCTACTGCCAGCTAGAGGCTAAATCCTGCCACATCCTGTTGG 2017

Qy   1975 AGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCC 2034
Db   2018 ACCAGCTGGGTACCTACGTGTTCACGGGCGAGTCCTACTCCCGCTCCGCAGTCAAGCGGC 2077

Qy   2035 TCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCT 2094
Db   2078 TCCAGCTAGCCATCTTCGCCCCAGCCCTCTGCACCTCCCTGGAGTATAGTCTCAGGGTCT 2137

Qy   2095 ACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGG 2154
Db   2138 ACTGTCTGGAGGACACTCCTGCAGCACTGAAGGAGGTCCTAGAGCTGGAGAGGACTCTGG 2197

Qy   2155 GGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAGTTACCACAACCTGC 2214
Db   2198 GTGGCTACTTGGTGGAGGAGCCCAAGACTTTGCTCTTTAAGGACAGTTACCACAACCTAC 2257

Qy   2215 GCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACC 2274
Db   2258 G-CTCTCCCTCCATGACATCCCCCATGCCCACTGGAGGAGCAAACTACTGGCCAAGTACC 2316

Qy   2275 AGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTTGCACTGCACCTTCA 2334
Db   2317 AGGAGATTCCCTTCTACCATGTGTGGAACGGCAGCCAGAAAGCCCTGCACTGCACTTTCA 2376
```

```
Qy       2335 CCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGG 2394
Db       2377 CCCTGGAGAGACATAGCCTAGCCTCCACTGAGTTCACCTGTAAGGTCTGCGTGCGGCAGG 2436

Qy       2395 TGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCAC---CAAGGACACAAGGT 2451
Db       2437 TAGAAGGGGAAGGCCAGATTTTCCAGCTGCACACCACGCTGGCTGAGACGCCTGCTGGCT 2496

Qy       2452 TTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTG 2511
Db       2497 CCCTGGATGCACTCTGCTCTGCCCCTGGCAATGCTGCCACCACACAGCTGGGACCCTATG 2556

Qy       2512 CCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTGGACCCACCCTGTA 2571
Db       2557 CCTTCAAGATACCACTGTCCATCCGCCAGAAGATCTGCAACAGCCTGGACGCCCCCAACT 2616

Qy       2572 GGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGACAGCCATCTCAGCT 2631
Db       2617 CACGGGGCAATGACTGGCGGCTGTTGGCACAGAAGCTCTCCATGGACCGGTACCTGAACT 2676

Qy       2632 TCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGGGAGGCGCGGCACT 2691
Db       2677 ACTTCGCCACCAAAGCTAGTCCCACAGGCGTGATCTTAGACCTCTGGGAAGCTCGGCAGC 2736

Qy       2692 TCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGGCCAGCCAGACG 2751
Db       2737 AGGATGATGGGGACCTCAACAGCCTGGCCAGTGCCTTGGAGGAGATGGGCAAGAGTGAGA 2796

Qy       2752 CTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2786
Db       2797 TGCTGGTAGCCATGACCACTGATGGCGATTGCTGA 2831
```

```
RESULT 8
US-09-833-381-1807
; Sequence 1807, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
;  APPLICANT: Robison, Keith E.
;  TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
;  FILE REFERENCE: 5800-119
;  CURRENT APPLICATION NUMBER: US/09/833,381
;  CURRENT FILING DATE:  2001-04-11
;  PRIOR APPLICATION NUMBER: 09/516,448
;  PRIOR FILING DATE: 2000-02-29
;  NUMBER OF SEQ ID NOS: 2050
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1807
;   LENGTH: 1605
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(1605)
;   OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1807

  Query Match             14.5%;  Score 419;  DB 4;  Length 1605;
  Best Local Similarity   62.4%;  Pred. No. 6.6e-82;
  Matches  737;  Conservative    0;  Mismatches  435;  Indels   10;  Gaps    5;

Qy       1556 CCTGCCCCGAGGCACCAGCAACATGACCTATGGACCTTCAACTTCCTCGGGGGCCGGCT 1615
Db        421 CCTGCCCCGAGACCCAGGGAGCAGCGTCAGCGGCACCTTTGGCTGCCTGGGTGGAGGCT 480

Qy       1616 GATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCCAGATGCCATACCCCGAGGGAA 1675
Db        481 CANCATCCCCGGCACAGGGGTCAGCTTGCTGGTGCCCAATGGAGCCATTCCCCAGGGCAA 540

Qy       1676 GATCTATGA--GATCTACCTCACGCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCT 1733
Db        541 GTTCTACGAAGATGTATTCTACTCATCAACAAGGCAGAAAGTACCCTCCCGCTTTCAGAA 600

Qy       1734 GGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCTGGCGTCCTGCTCACC 1793
Db        601 GGGACCCANACAGTATTGAGCCCCTCGGTGACCTGTGGACCCACAGGCCTCCTGCTGTGC 660

Qy       1794 CGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTG 1853
Db        661 CGCCCCGTCATCCTCACCATGCCCCACTGTGCCGAAGTCAGTGCCCGTGACTGGATCTTT 720

Qy       1854 CGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGCAGGATGTGCTGCACCTGGGCGAG 1913
Db        721 CAGCTCAAGACCCAGGCCCACCAGGGCCACTGGG---ANGAGGTGGTGACCCTGGATGAG 777

Qy       1914 GAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACC 1973
Db        778 GAGACCCTGAACACACCCTGCTACTGCCAGCTGGAGCCCAGGGCCTGTCACATCCTGCTG 837

Qy       1974 GAGCAGCTGGGCCGCT-TTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCG 2032
Db        838 GACCAGCTGGGCACCTACCGTGTTCACGGGCGAGTCCTATTCCCGCTCAGCAGTCAAGCG 897

Qy       2033 CCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAGTACAACATCCGGGT 2092
Db        898 GCTCCAGCTGGCCGTCTTCGCCCCCGCCCTCTGCACCTCCCTGGAGTACAGCCTCCGGGT 957

Qy       2093 CTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCT 2152
Db        958 CTACTGCCTGGAGGACACGCCTGTAGCACTGAAGGAGGTGCTGGAGCTGGAGCGGACTCT 1017

Qy       2153 GGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAGTTACCACAACCT 2212
Db       1018 GGGCGGATACTTGGTGGAGGAGCCGAAACCGCTAATGTTCAAGGACAGTTACCACAACCT 1077

Qy       2213 GCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTA 2272
Db       1078 GCGCCTCTCCCTCCATGACCTCCCCCATGCCCATTGGAGGAGCAAGCTGCTGGCCAAATA 1137

Qy       2273 CCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTTGCACTGCACCTT 2332
Db       1138 CCAGGAGATCCCCTTCTATCACATTTGGAGTGGCAGCCAGAAGGCCCTCCACTGCACTTT 1197

Qy       2333 CACCCTGGAGCGTGTCAGCCCCAG-CACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGC 2391
Db       1198 CACCCTGGAGAGGCACAGGCTTGGCCTCCACAGAGCTCACCTGCAAGATCTGCGTGCGGC 1257

Qy       2392 AGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACCAAG---GACACAA 2448
Db       1258 AAGTGGAAGGGGAGGGCCAGATATTCCAGCTGCATACCACTCTGGCAGAGACACCTGCTG 1317

Qy       2449 GGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTGGTGGGCCCCA 2508
Db       1318 GCTCCCTGGACACTCTCTGCTCTGCCCCTGGCAGCACTGTCACCACCCAGCTGGGACCTT 1377

Qy       2509 GTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTGGACCCACCCT 2568
Db       1378 ATGCCTTCAAGATCCCACTGTCCATCCGCCAGAAGATATGCAACAGCCTAGATGCCCCCA 1437

Qy       2569 GTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGACAGCCATCTCA 2628
Db       1438 ACTCACGGGGCAATGACTGGCGGATGTTAGCACAGAAGCTCTCTATGGACCGGTACCTGA 1497

Qy       2629 GCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGGGAGGCGCGGC 2688
Db       1498 ATTACTTTGCCACCAAAGCGAGCCCCACGGNTGTGATCCTGGACCTCTGGGAAGCTCTGC 1557

Qy       2689 ACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGG 2730
Db       1558 AGCAGGACGATGGGGACCTCAACAGCCTGNCGAGTGCCTTGG 1599
```

```
RESULT 9
US-08-253-155A-17/c
; Sequence 17, Application US/08253155A
```

```
; Patent No. 5691147
;  GENERAL INFORMATION:
;    APPLICANT:  Gyuris, Jeno
;    APPLICANT:  Draetta, Giulio
;    TITLE OF INVENTION:  CDK4 Binding Proteins
;    NUMBER OF SEQUENCES:  95
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD
;      STREET:  60 State Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII(text)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/253,155A
;      FILING DATE:  02-JUN-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Vincent, Matthew P.
;      REGISTRATION NUMBER:  36,709
;      REFERENCE/DOCKET NUMBER:  MII-028
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 227-7400
;      TELEFAX:  (617) 227-5941
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  771 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-253-155A-17

  Query Match             8.9%;  Score 256;  DB 1;  Length 771;
  Best Local Similarity   67.2%;  Pred. No. 1.3e-46;
  Matches  406;  Conservative    1;  Mismatches  191;  Indels    6;  Gaps    3;

Qy       1741 AGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAG 1800

Db        738 AGACAGTATTGAGCCCCTCGGTGACCTGTGGACCCACAGGCCTCCTGCTGTGCCGCCCCG 679

Qy       1801 TCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCA 1860

Db        678 TCATCCTCACCATGCCCCACTGTGCCGAAGTCAGTGCCCGTGACTGGATCTTTCAGCTCA 619

Qy       1861 AAAAGCAGTCGTGCGAGGGCAGCTGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGC 1920

Db        618 AGACCCAGGCCCACCAGGGCCACTGGG---AGGAGGTGGTGACCCTGGATGAGGAGACCC 562

Qy       1921 CCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGC 1980

Db        561 TGAACACACCCTGCTACTGCCAGCTGGAGCCCAGGGCCTGTCACATCCTGCTGGACCAGC 502

Qy       1981 TGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGC 2040

Db        501 TGGGCACCTACGTGTTCACGGGCGAGTCCTATTCCCGCTCAGCAGTCAAGCGGSTCCAGC 442

Qy       2041 TGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCC 2100

Db        441 TGGCCGTCTTCGCCCCCGCCCTCTGCACCTCCCTGGAGTACAGACTCCGGGTCTACTGTC 382

Qy       2101 TGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGAC 2160

Db        381 TGGAGGACAGGCCTGTAGCACTGAAGGAGGTGCTGGAGCTGGAGCGGACTCTGGGCGGAT 322

Qy       2161 AGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTAT 2220

Db        321 ACTTGGTGGAGGAGCCGAAACCGCTAATGTTCAAGGACAGTTAACACAACCTGCG-CTCT 263

Qy       2221 CCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGA 2280

Db        262 TCCTCCATGACCTCCCCCATGCCCATTGGAGGAGCAAGCTGCTGGCCAAATACCAGGAGA 203

Qy       2281 TCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGG 2340

Db        202 TCCCCTTCTATCACATTTGGAGTGGAGCCAGAAGG--CCTCCACTGCACTTTCACCCTGG 145

Qy       2341 AGCG 2344

Db        144 AGAG 141


RESULT 10
US-08-808-982-4
; Sequence 4, Application US/08808982
; Patent No. 5939271
;  GENERAL INFORMATION:
;    APPLICANT:  Tessier-Lavigne, Marc
;    APPLICANT:  Leonardo, E. David
;    APPLICANT:  Hink, Lindsay
;    APPLICANT:  Masu, Masayuki
;    APPLICANT:  Kazuko, Keino-Masu
;    TITLE OF INVENTION:  Netrin Receptors
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SCIENCE & TECHNOLOGY LAW GROUP
;      STREET:  268 BUSH STREET, SUITE 3200
;      CITY:  SAN FRANCISCO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/808,982
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  OSMAN, RICHARD A
;      REGISTRATION NUMBER:  36,627
;      REFERENCE/DOCKET NUMBER:  UC96-217
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 343-4341
;      TELEFAX:  (415) 343-4342
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  305 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-808-982-4

  Query Match             4.1%;  Score 119;  DB 2;  Length 305;
  Best Local Similarity   68.1%;  Pred. No. 5.6e-17;
  Matches  209;  Conservative    0;  Mismatches   95;  Indels    3;  Gaps    3;

Qy       1906 TGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTGCTACG 1965

Db          1 TGGATGAGGAGACCCTGAACACACCCTGCTACTG-CAGCTGGAGCCCAGGGCCTG-TACA 58

Qy       1966 TCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCG 2025

Db         59 TCCTGCTGGACCAGCTGGGCACCTACGTTTTCACGGGCGAGTCCTATTCCCGCTCAGCAG 118

Qy       2026 CCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAGTACAACA 2085

Db        119 TCAAGCGGCTCCAGCTGGCCGT-TTCGCCCCCGCCCTCTGCACCTCCCTGGAGTACAGCC 177
```

```
Qy       2086 TCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTGGAGA 2145
              ||||||||||||||||| | ||||| |   |  ||||| |||||||||  || ||||||||
Db        178 TCCGGGTCTACTGCCTGGAGGACACGCCTGTAGCACTGAAGGAGGTGCTGGAGCTGGAGC 237

Qy       2146 AGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAGTTACC 2205
                |   ||||| ||| |  || |  |||||||        ||    |||||||||||||||
Db        238 GGACTCTGGGCGGATACTTGGTGGAGGAGCCGAAACCGCTAATGTTCAAGGACAGTTACC 297

Qy       2206 ACAACCT 2212
              |||||||
Db        298 ACAACCT 304
```

RESULT 11
US-09-306-902A-4

```
; Sequence 4, Application US/09306902A
; Patent No. 6277585
;  GENERAL INFORMATION:
;       APPLICANT: Tessier-Lavigne, Marc
;                  Leonardo, E. David
;                  Hink, Lindsay
;                  Masu, Masayuki
;                  Kazuko, Keino-Masu
;       TITLE OF INVENTION: Netrin Receptors
;       NUMBER OF SEQUENCES: 9
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
;            STREET: 268 BUSH STREET, SUITE 3200
;            CITY: SAN FRANCISCO
;            STATE: CALIFORNIA
;            COUNTRY: USA
;            ZIP: 94104
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/306,902A
;            FILING DATE: 07-May-1999
;            CLASSIFICATION: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: OSMAN, RICHARD A
;            REGISTRATION NUMBER: 36,627
;            REFERENCE/DOCKET NUMBER: UC96-217
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (415) 343-4341
;            TELEFAX: (415) 343-4342
;  INFORMATION FOR SEQ ID NO: 4:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 305 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: double
;            TOPOLOGY: linear
;       MOLECULE TYPE: cDNA
;       SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-306-902A-4

  Query Match             4.1%;  Score 119;  DB 3;  Length 305;
  Best Local Similarity   68.1%;  Pred. No. 5.6e-17;
  Matches  209;  Conservative    0;  Mismatches   95;  Indels    3;  Gaps    3;

Qy       1906 TGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTGCTACG 1965
              |||  |||||| | |     |    | || |||||| ||||||||| |||| ||||| |||
Db          1 TGGATGAGGAGACCCTGAACACACCCTGCTACTG-CAGCTGGAGCCCAGGGCCTG-TACA 58

Qy       1966 TCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCG 2025
              || |     || |||||||||  ||  |    |     ||| ||| ||     |    | || |
Db         59 TCCTGCTGGACCAGCTGGGCACCTACGTTTTCACGGGCGAGTCCTATTCCCGCTCAGCAG 118

Qy       2026 CCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAGTACAACA 2085
              |||||| ||| |||||      | || || || |      ||||||||||| ||||||| |
```



```
Db        119 TCAAGCGGCTCCAGCTGGCCGT-TTCGCCCCCGCCCTCTGCACCTCCCTGGAGTACAGCC 177

Qy       2086 TCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTGGAGA 2145
              ||||||||||||||||| | ||||| |   |  ||||| |||||||||  || ||||||||
Db        178 TCCGGGTCTACTGCCTGGAGGACACGCCTGTAGCACTGAAGGAGGTGCTGGAGCTGGAGC 237

Qy       2146 AGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAGTTACC 2205
                |   ||||| ||| |  || |  |||||||        ||    |||||||||||||||
Db        238 GGACTCTGGGCGGATACTTGGTGGAGGAGCCGAAACCGCTAATGTTCAAGGACAGTTACC 297

Qy       2206 ACAACCT 2212
              |||||||
Db        298 ACAACCT 304
```



RESULT 12
US-08-483-533-38

```
; Sequence 38, Application US/08483533
; Patent No. 6172047
;  GENERAL INFORMATION:
;    APPLICANT:  Roizman, Bernard
;    APPLICANT:  Chou, Joany
;    TITLE OF INVENTION:  Method for Treating Tumorigenic
;    TITLE OF INVENTION:  Diseases
;    NUMBER OF SEQUENCES:  43
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/483,533
;      FILING DATE:  07-MAR-95
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/419,853
;      FILING DATE:  11-APR-95
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/861,233
;      FILING DATE:  31-MAR-92
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Zeller, James P.
;      REGISTRATION NUMBER:  28,491
;      REFERENCE/DOCKET NUMBER:  28097/32742
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  38:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1280 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-483-533-38

  Query Match             1.8%;  Score 51.8;  DB 3;  Length 1280;
  Best Local Similarity   51.0%;  Pred. No. 0.034;
  Matches  122;  Conservative    0;  Mismatches  117;  Indels    0;  Gaps    0;

Qy          6 GGGCTCCGGGCTGAGGCGCTAAAGCCGCCCTCCCGCCCGCGGGGCCCCGCGCCCGGCCCG 65
              | |  | |||  | |||||  ||||| ||| |    |||    |  ||||||| ||   | ||
Db        555 GCGGGCGGGGAGGGGGCGCCGAAGCCCCCCGCGACCCCCGCGACCCCCGCGACCCCCGCG 614

Qy         66 CCCGCCTGCCCGCCCGCGGCCATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGC 125
```

```
Db        615 ACCCCCGCGACCCCCGCGACCCCCGCGACCCCCGCGCGGGTGCGCTTCTCGCCCCACGTC 674

Qy        126 ATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAAC 185

Db        675 CGGGTGCGCCACCTGGTGGTCTGGGCCTCGGCCGCCCGCCTGGCGCGCCGCGGCTCGTGG 734

Qy        186 CCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGT 244

Db        735 GCCCGCGAGCGGGCCGACCGGGCTCGGTTCCGGCGCCGGGTGGCGGAGGCCGAGGCGGT 793
```

RESULT 13
US-09-283-471A-38
; Sequence 38, Application US/09283471A
; Patent No. 6340673
;  GENERAL INFORMATION:
;    APPLICANT:  Roizman, Bernard
;    APPLICANT:  Chou, Joany
;    TITLE OF INVENTION:  Method For Treating Tumorigenic Diseases
;    NUMBER OF SEQUENCES:  43
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/283,471A
;      FILING DATE:  04-APR-1999
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/861,233
;      FILING DATE:  31-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/419,853
;      FILING DATE:  11-APR-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/483,533
;      FILING DATE:  07-JUN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Zeller, James P.
;      REGISTRATION NUMBER:  28,491
;      REFERENCE/DOCKET NUMBER:  27373/32742A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  38:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1280 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-09-283-471A-38

```
  Query Match             1.8%;  Score 51.8;  DB 4;  Length 1280;
  Best Local Similarity   51.0%;  Pred. No. 0.034;
  Matches  122;  Conservative    0;  Mismatches  117;  Indels    0;  Gaps    0;

Qy          6 GGGCTCCGGGCTGAGGCGCTAAAGCCGCCCTCCCGCCCGCGGGGCCCCGCGCCCGGCCCG 65

Db        555 GCGGGCGGGGAGGGGGCGCCGAAGCCCCCCGCGACCCCCGCGACCCCCGCGACCCCCGCG 614

Qy         66 CCCGCCTGCCCGCCCGCGGCCATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGC 125
```


```
Db        615 ACCCCCGCGACCCCCGCGACCCCCGCGACCCCCGCGCGGGTGCGCTTCTCGCCCCACGTC 674

Qy        126 ATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAAC 185

Db        675 CGGGTGCGCCACCTGGTGGTCTGGGCCTCGGCCGCCCGCCTGGCGCGCCGCGGCTCGTGG 734

Qy        186 CCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGT 244

Db        735 GCCCGCGAGCGGGCCGACCGGGCTCGGTTCCGGCGCCGGGTGGCGGAGGCCGAGGCGGT 793
```


RESULT 14
US-09-252-991A-7947
; Sequence 7947, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 7947
;   LENGTH: 699
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7947

```
  Query Match             1.8%;  Score 51.2;  DB 4;  Length 699;
  Best Local Similarity   47.5%;  Pred. No. 0.038;
  Matches  152;  Conservative    0;  Mismatches  168;  Indels    0;  Gaps    0;

Qy       1876 AGGGCAGCTGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACT 1935

Db        168 ACGACGTCTACGACAGCAACGAGGTGCTCGAGGCCAAGCTGCTGTCCGGGCATTCCGGCT 227

Qy       1936 ACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCC 1995

Db        228 ACGACCTGGTGGTGCCCAGCGACAGCTTCCTGCCCAACTACCTGAAGGCCGAGGTGTTCC 287

Qy       1996 TGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGC 2055

Db        288 AGCCGCTGGACAAGAGCAAGCTGCCGAACTGGAAGAACCTCAACCCGGCCCTGCTCAAGG 347

Qy       2056 CGGTGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACG 2115

Db        348 TGCTCGCCGGCAAGGACCCCGGCAACCGCTACGTGATGCCCTACATGTGGGGCACCAACG 407

Qy       2116 ATGCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGC 2175

Db        408 GCATCGCCTACAACCTCGACAAGGTCCGCGCGGTGCTCGGCGACGATGCGCCGCTGGACT 467

Qy       2176 CACGGGTCCTGCACTTCAAG 2195

Db        468 CCTGGGACCTGGTGTTCAAG 487
```

RESULT 15
US-09-252-991A-7809
; Sequence 7809, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18

```
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 7809
;   LENGTH: 1302
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7809

  Query Match              1.8%;  Score 51.2;  DB 4;  Length 1302;
  Best Local Similarity   47.5%;  Pred. No. 0.046;
  Matches  152;  Conservative    0;  Mismatches  168;  Indels    0;  Gaps    0;

Qy       1876 AGGGCAGCTGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACT 1935

Db        368 ACGACGTCTACGACAGCAACGAGGTGCTCGAGGCCAAGCTGCTGTCCGGGCATTCCGGCT 427

Qy       1936 ACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCC 1995

Db        428 ACGACCTGGTGGTGCCCAGCGACAGCTTCCTGCCCAACTACCTGAAGGCCGAGGTGTTCC 487

Qy       1996 TGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGC 2055

Db        488 AGCCGCTGGACAAGAGCAAGCTGCCGAACTGGAAGAACCTCAACCCGGCCCTGCTCAAGG 547

Qy       2056 CGGTGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACG 2115

Db        548 TGCTCGCCGGCAAGGACCCCGGCAACCGCTACGTGATGCCCTACATGTGGGGCACCAACG 607

Qy       2116 ATGCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGC 2175

Db        608 GCATCGCCTACAACCTCGACAAGGTCCGCGCGGTGCTCGGCGACGATGCGCCGCTGGACT 667

Qy       2176 CACGGGTCCTGCACTTCAAG 2195

Db        668 CCTGGGACCTGGTGTTCAAG 687


Search completed: October  6, 2004, 17:53:37
Job time : 205 secs
```

```
OM nucleic - nucleic search, using sw model

Run on:         October 6, 2004, 12:26:42 ; Search time 1334 Seconds
                                            (without alignments)
                                            10948.592 Million cell updates/sec

Title:          US-09-970-944-1
Perfect score:  2881
Sequence:       1 agctggggctccgggctgag..........ccttccccacaccggggaga 2881

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       3340653 seqs, 2534783454 residues

Total number of hits satisfying chosen parameters:      6681306

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Published_Applications_NA:*
                 1:  /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
                 2:  /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
                 3:  /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
                 4:  /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
                 5:  /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
                 6:  /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
                 7:  /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
                 8:  /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
                 9:  /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
                 10:  /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
                 11:  /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
                 12:  /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
                 13:  /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
                 14:  /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
                 15:  /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
                 16:  /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
                 17:  /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
                 18:  /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
                 19:  /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 2881 | 100.0 | 2881 | 11 | US-09-970-944-1 | Sequence 1, Appli |
| 2 | 2676.4 | 92.9 | 2752 | 13 | US-09-918-779-1 | Sequence 1, Appli |
| 3 | 2676.4 | 92.9 | 2752 | 17 | US-10-624-932-1 | Sequence 1, Appli |
| 4 | 2402.6 | 83.4 | 3580 | 17 | US-10-311-623-13 | Sequence 13, Appl |
| 5 | 2235.8 | 77.6 | 3014 | 10 | US-09-933-261-1 | Sequence 1, Appli |
| 6 | 2235.8 | 77.6 | 3014 | 15 | US-10-256-702-1 | Sequence 1, Appli |
| 7 | 2189.4 | 76.0 | 2697 | 15 | US-10-240-154-15 | Sequence 15, Appl |
| 8 | 1570.4 | 54.5 | 1787 | 10 | US-09-933-261-2 | Sequence 2, Appli |
| 9 | 1570.4 | 54.5 | 1787 | 15 | US-10-256-702-2 | Sequence 2, Appli |
| 10 | 1193.6 | 41.4 | 1321 | 13 | US-10-296-115-365 | Sequence 365, App |
| 11 | 1147.8 | 39.8 | 1282 | 9 | US-09-833-381-1806 | Sequence 1806, Ap |
| 12 | 889 | 30.9 | 2860 | 13 | US-10-087-684-1 | Sequence 1, Appli |
| 13 | 889 | 30.9 | 2860 | 13 | US-10-218-779-1 | Sequence 1, Appli |
| 14 | 887.4 | 30.8 | 2860 | 13 | US-10-087-684-3 | Sequence 3, Appli |
| 15 | 887.4 | 30.8 | 2860 | 13 | US-10-218-779-3 | Sequence 3, Appli |
| 16 | 875.8 | 30.4 | 2995 | 13 | US-09-972-211-55 | Sequence 55, Appl |
| 17 | 875.8 | 30.4 | 2995 | 13 | US-10-096-625-55 | Sequence 55, Appl |
| 18 | 866.4 | 30.1 | 2895 | 13 | US-10-037-417-37 | Sequence 37, Appl |
| 19 | 855.2 | 29.7 | 3485 | 9 | US-09-816-828-18 | Sequence 18, Appl |
| 20 | 853.6 | 29.6 | 3884 | 13 | US-10-147-493-145 | Sequence 145, App |
| 21 | 853.6 | 29.6 | 3884 | 13 | US-10-145-127-145 | Sequence 145, App |
| 22 | 853.6 | 29.6 | 3884 | 13 | US-10-160-503-145 | Sequence 145, App |
| 23 | 853.6 | 29.6 | 3884 | 13 | US-10-143-118-145 | Sequence 145, App |
| 24 | 853.6 | 29.6 | 3884 | 13 | US-10-144-993-145 | Sequence 145, App |
| 25 | 853.6 | 29.6 | 3884 | 13 | US-10-158-787-145 | Sequence 145, App |
| 26 | 853.6 | 29.6 | 3884 | 13 | US-10-140-024-145 | Sequence 145, App |
| 27 | 853.6 | 29.6 | 3884 | 13 | US-10-140-808-145 | Sequence 145, App |
| 28 | 853.6 | 29.6 | 3884 | 13 | US-10-152-405-145 | Sequence 145, App |
| 29 | 853.6 | 29.6 | 3884 | 13 | US-10-127-852A-145 | Sequence 145, App |
| 30 | 853.6 | 29.6 | 3884 | 13 | US-10-127-900A-145 | Sequence 145, App |
| 31 | 853.6 | 29.6 | 3884 | 13 | US-10-128-685A-145 | Sequence 145, App |
| 32 | 853.6 | 29.6 | 3884 | 13 | US-10-131-820A-145 | Sequence 145, App |
| 33 | 853.6 | 29.6 | 3884 | 13 | US-10-142-886-145 | Sequence 145, App |
| 34 | 853.6 | 29.6 | 3884 | 13 | US-10-146-728-145 | Sequence 145, App |
| 35 | 853.6 | 29.6 | 3884 | 13 | US-10-146-786-145 | Sequence 145, App |
| 36 | 853.6 | 29.6 | 3884 | 13 | US-10-147-499-145 | Sequence 145, App |
| 37 | 853.6 | 29.6 | 3884 | 13 | US-10-157-798-145 | Sequence 145, App |
| 38 | 853.6 | 29.6 | 3884 | 15 | US-10-028-072-145 | Sequence 145, App |
| 39 | 853.6 | 29.6 | 3884 | 15 | US-10-121-049-145 | Sequence 145, App |
| 40 | 853.6 | 29.6 | 3884 | 15 | US-10-123-904-145 | Sequence 145, App |
| 41 | 853.6 | 29.6 | 3884 | 15 | US-10-140-470-145 | Sequence 145, App |
| 42 | 853.6 | 29.6 | 3884 | 15 | US-10-175-746-145 | Sequence 145, App |
| 43 | 853.6 | 29.6 | 3884 | 15 | US-10-176-918-145 | Sequence 145, App |
| 44 | 853.6 | 29.6 | 3884 | 15 | US-10-176-921-145 | Sequence 145, App |
| 45 | 853.6 | 29.6 | 3884 | 15 | US-10-137-865-145 | Sequence 145, App |

ALIGNMENTS

```
RESULT 1
US-09-970-944-1
; Sequence 1, Application US/09970944
; Publication No. US20030204052A1
; GENERAL INFORMATION:
;  APPLICANT: Herrman, John L
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Shimkets, Richard A
;  TITLE OF INVENTION: No. US20030204052A1el Proteins and Nucleic Acids Encoding Same anc
;  TITLE OF INVENTION:  Antibodies Directed Against these Proteins
;  FILE REFERENCE: 21402-138
;  CURRENT APPLICATION NUMBER: US/09/970,944
;  CURRENT FILING DATE:  2002-05-02
;  PRIOR APPLICATION NUMBER: 60/237,862
;  PRIOR FILING DATE: 2000-10-04
;  NUMBER OF SEQ ID NOS: 62
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 2881
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-970-944-1

  Query Match             100.0%;  Score 2881;  DB 11;  Length 2881;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 2881;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGCTGGGGCTCCGGGCTGAGGCGCTAAAGCCGCCCTCCCGCCCGCGGGGCCCCGCGCCCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGCTGGGGCTCCGGGCTGAGGCGCTAAAGCCGCCCTCCCGCCCGCGGGGCCCCGCGCCCG 60

Qy         61 GCCCGCCCGCCTGCCCGCCCGCGGCCATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCCCGCCCGCCTGCCCGCCCGCGGCCATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCC 120

Qy        121 TGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      121 TGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGG 180

Qy      181 CCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGG 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 CCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGG 240

Qy      241 ATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGC 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 ATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGC 300

Qy      301 AGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCA 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 AGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCA 360

Qy      361 GCACAGACGGGAGCAGTGGTGAGCCGACCATGGAGGTCCGCATTAATGTCTCAAGGCAGC 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 GCACAGACGGGAGCAGTGGTGAGCCGACCATGGAGGTCCGCATTAATGTCTCAAGGCAGC 420

Qy      421 AGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCT 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 AGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCT 480

Qy      481 CGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCG 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 CGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCG 540

Qy      541 AGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTC 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      541 AGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTC 600

Qy      601 CACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACC 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      601 CACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACC 660

Qy      661 CGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCC 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      661 CGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCC 720

Qy      721 GCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCA 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      721 GCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCA 780

Qy      781 GCGCCTCCGCTGCTGTCATCGTCTACGTGAACGGTGGGTGGTCGACGTGGACCGAGTGGT 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      781 GCGCCTCCGCTGCTGTCATCGTCTACGTGAACGGTGGGTGGTCGACGTGGACCGAGTGGT 840

Qy      841 CCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAACC 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      841 CCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAACC 900

Qy      901 CGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAATGTCCATGACCGCACCGTCT 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      901 CGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAATGTCCATGACCGCACCGTCT 960

Qy      961 CCTCTCTGCTTGTCTCTGTGGACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTG 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      961 CCTCTCTGCTTGTCTCTGTGGACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTG 1020

Qy     1021 GGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAG 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1021 GGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAG 1080

Qy     1081 GGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGTACCAGTGACCTCTGTGTAC 1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1081 GGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGTACCAGTGACCTCTGTGTAC 1140

Qy     1141 ACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCT 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1141 ACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCT 1200

Qy     1201 GCCTGGTCCTGCTGCTGCTTGTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGG 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1201 GCCTGGTCCTGCTGCTGCTTGTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGG 1260

Qy     1261 ACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGC 1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1261 ACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGC 1320

Qy     1321 CCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCA 1380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1321 CCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCA 1380

Qy     1381 CCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCA 1440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1381 CCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCA 1440

Qy     1441 ATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCA 1500
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1441 ATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCA 1500

Qy     1501 CCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGC 1560
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1501 CCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGC 1560

Qy     1561 CCCGAGGCACCAGCAACATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGA 1620
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1561 CCCGAGGCACCAGCAACATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGA 1620

Qy     1621 TCCCTAATACAGGTATCAGCCTCCTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCT 1680
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1621 TCCCTAATACAGGTATCAGCCTCCTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCT 1680

Qy     1681 ATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTC 1740
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1681 ATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTC 1740

Qy     1741 AGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAG 1800
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1741 AGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAG 1800

Qy     1801 TCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCA 1860
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1801 TCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCA 1860

Qy     1861 AAAAGCAGTCGTGCGAGGGCAGCTGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGC 1920
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1861 AAAAGCAGTCGTGCGAGGGCAGCTGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGC 1920

Qy     1921 CCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGC 1980
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1921 CCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGC 1980

Qy     1981 TGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGC 2040
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1981 TGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGC 2040

Qy     2041 TGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCC 2100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2041 TGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCC 2100

Qy     2101 TGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGAC 2160
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2101 TGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGAC 2160

Qy     2161 AGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTAT 2220
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2161 AGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTAT 2220

Qy     2221 CCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGA 2280
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2221 CCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGA 2280

Qy     2281 TCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGG 2340
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2281 TCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGG 2340
```

```
Qy       2341 AGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGG 2400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 AGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGG 2400

Qy       2401 GCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGC 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 GCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGC 2460

Qy       2461 TGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGA 2520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 TGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGA 2520

Qy       2521 TCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTG 2580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 TCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTG 2580

Qy       2581 CCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCT 2640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 CCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCT 2640

Qy       2641 CCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACG 2700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 CCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACG 2700

Qy       2701 GCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCT 2760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 GCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCT 2760

Qy       2761 TCACAGTGTCGGAGGCTGAGTGCTGAGGCCGGCCAGGCCCGACACCTACACTCTCACCAG 2820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 TCACAGTGTCGGAGGCTGAGTGCTGAGGCCGGCCAGGCCCGACACCTACACTCTCACCAG 2820

Qy       2821 CTTTGGCACCCACCAAGGACAGGCAGAAGCCGGACAGGGGCCCTTCCCCACACCGGGGAG 2880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 CTTTGGCACCCACCAAGGACAGGCAGAAGCCGGACAGGGGCCCTTCCCCACACCGGGGAG 2880

Qy       2881 A 2881
              |
Db       2881 A 2881
```

```
RESULT 2
US-09-918-779-1
; Sequence 1, Application US/09918779
; Publication No. US20030064369A1
; GENERAL INFORMATION:
;  APPLICANT: Taupier, Raymond
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Spaderna, Steven
;  APPLICANT:  Shimkets, Richard
;  APPLICANT:  Zerhusen, Bryan
;  APPLICANT:  Spytek, Kimberly
;  APPLICANT:  Shenoy, Suresh
;  APPLICANT:  Li, Li
;  APPLICANT:  Gusev, Vladimir
;  APPLICANT:  Grosse, William
;  APPLICANT:  Alsobrook, John
;  APPLICANT:  Lepley, Denise
;  APPLICANT:  Burgess, Catherine
;  APPLICANT:  Gerlach, Valerie
;  APPLICANT:  Ellerman, Karen
;  APPLICANT:  MacDougall, John
;  APPLICANT:  Stone, David
;  APPLICANT:  Smithson, Glennda
;  TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
;  FILE REFERENCE: 21402-074 US
;  CURRENT APPLICATION NUMBER: US/09/918,779
;  CURRENT FILING DATE:  2001-07-30
;  PRIOR APPLICATION NUMBER: 60/221,409
;  PRIOR FILING DATE: 2000-07-28
;  PRIOR APPLICATION NUMBER: 60/222,840
;  PRIOR FILING DATE: 2000-08-04
;  PRIOR APPLICATION NUMBER: 60/223,752
;  PRIOR FILING DATE: 2000-08-08
;  PRIOR APPLICATION NUMBER: 60/223,762
;  PRIOR FILING DATE: 2000-08-08
;  PRIOR APPLICATION NUMBER: 60/223,770
;  PRIOR FILING DATE: 2000-08-08
;  PRIOR APPLICATION NUMBER: 60/223,769
;  PRIOR FILING DATE: 2000-08-08
;  PRIOR APPLICATION NUMBER: 60/225,146
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/225,392
;  PRIOR FILING DATE: 2000-08-15
;  PRIOR APPLICATION NUMBER: 60/225,470
;  PRIOR FILING DATE: 2000-08-15
;  PRIOR APPLICATION NUMBER: 60/225,697
;  PRIOR FILING DATE: 2000-08-16
;  PRIOR APPLICATION NUMBER: 60/263,662
;  PRIOR FILING DATE: 2001-02-01
;  PRIOR APPLICATION NUMBER: 60/281,645
;  PRIOR FILING DATE: 2001-04-05
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 2752
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-918-779-1

  Query Match             92.9%;  Score 2676.4;  DB 13;  Length 2752;
  Best Local Similarity   98.9%;  Pred. No. 0;
  Matches 2728;  Conservative    0;  Mismatches   21;  Indels    9;  Gaps    3;

Qy         42 CCGCGGGGCCCCGCGCCCGGCCCGCCCGCCTGCCCGCCCGCGGCCATGGCCGTCCGGCCC 101
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCGCGGGGCCCCGCGCCCGGCCCGCCCGCCTGCCCGCCCGCGGCCATGGCCGTCCGGCCC 60

Qy        102 GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC 161
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC 120

Qy        162 CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC 221
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC 180

Qy        222 TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC 281
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC 240

Qy        282 AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG 341
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG 300

Qy        342 GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCGACCATGGAGGTCCGC 401
              ||||||||||||||||||||||||||||||||||||||   ||| |||||||||||||||
Db        301 GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCCACCATGGAGGTCCGC 360

Qy        402 ATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAG 461
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAG 420

Qy        462 TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC 521
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC 480

Qy        522 AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC 581
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC 540

Qy        582 ATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG 641
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG 600

Qy        642 AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC 701
```

```
Db    601 AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC 660

Qy    702 CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC 761
Db    661 CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC 720

Qy    762 ATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTACGTGAACGGTGGGTGG 821
Db    721 ATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTACGTGAACGGTGGGTGG 780

Qy    822 TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG 881
Db    781 TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG 840

Qy    882 AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT 941
Db    841 AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT 900

Qy    942 GTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGCAGCTGGAGCCCGTGG 1001
Db    901 GTCCAGAA---AACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGCTGGAGCCCGTGG 957

Qy   1002 AGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC 1061
Db    958 AGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC 1017

Qy   1062 CCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT 1121
Db   1018 CCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT 1077

Qy   1122 ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC 1181
Db   1078 ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC 1137

Qy   1182 CTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCATCCTCGTTTATTGC 1241
Db   1138 CTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCATCCTCGTTTATTGC 1197

Qy   1242 CGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC 1301
Db   1198 CGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC 1257

Qy   1302 CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG 1361
Db   1258 CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG 1317

Qy   1362 GACCTCAG---CACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC 1418
Db   1318 GACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC 1377

Qy   1419 AGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC 1478
Db   1378 AGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC 1437

Qy   1479 ACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC 1538
Db   1438 ACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC 1497

Qy   1539 CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC 1598
Db   1498 CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC 1557

Qy   1599 TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCCAGAT 1658
Db   1558 TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCCAGAT 1617

Qy   1659 GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG 1718
Db   1618 GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG 1677

Qy   1719 AGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT 1778
Db   1678 AGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT 1737

Qy   1779 GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT 1838
Db   1738 GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT 1797

Qy   1839 GACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGCAGGATGTG 1898
Db   1798 GACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGG---AGGATGTG 1854

Qy   1899 CTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC 1958
Db   1855 CTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC 1914

Qy   1959 TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG 2018
Db   1915 TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG 1974

Qy   2019 GCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAG 2078
Db   1975 GCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAG 2034

Qy   2079 TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG 2138
Db   2035 TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG 2094

Qy   2139 CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC 2198
Db   2095 CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC 2154

Qy   2199 AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAG 2258
Db   2155 AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAG 2214

Qy   2259 CTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC 2318
Db   2215 CTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC 2274

Qy   2319 TTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAG 2378
Db   2275 TTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAG 2334

Qy   2379 CTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACC 2438
Db   2335 CTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACC 2394

Qy   2439 AAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG 2498
Db   2395 AAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG 2454

Qy   2499 GTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTG 2558
Db   2455 GTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTG 2514

Qy   2559 GACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGAC 2618
Db   2515 GACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGAC 2574

Qy   2619 AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGG 2678
Db   2575 AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGG 2634

Qy   2679 GAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTG 2738
Db   2635 GAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTG 2694

Qy   2739 GGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGAGGCCGGCCAG 2796
Db   2695 GGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGAGGCCGGCCAG 2752
```

RESULT 3
US-10-624-932-1

```
; Sequence 1, Application US/10624932
; Publication No. US20040096877A1
; GENERAL INFORMATION:
;  APPLICANT: Taupier, Raymond
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Spaderna, Steven
;  APPLICANT:  Shimkets, Richard
;  APPLICANT:  Zerhusen, Bryan
;  APPLICANT:  Spytek, Kimberly
;  APPLICANT:  Shenoy, Suresh
;  APPLICANT:  Li, Li
;  APPLICANT:  Gusev, Vladimir
;  APPLICANT:  Grosse, William
;  APPLICANT:  Alsobrook, John
;  APPLICANT:  Lepley, Denise
;  APPLICANT:  Burgess, Catherine
;  APPLICANT:  Gerlach, Valerie
;  APPLICANT:  Ellerman, Karen
;  APPLICANT:  MacDougall, John
;  APPLICANT:  Stone, David
;  APPLICANT:  Smithson, Glennda
;  TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
;  FILE REFERENCE: 21402-074 US
;  CURRENT APPLICATION NUMBER: US/10/624,932
;  CURRENT FILING DATE:  2003-07-21
;  PRIOR APPLICATION NUMBER: 09/918,779
;  PRIOR FILING DATE: 2001-07-03
;  PRIOR APPLICATION NUMBER: 60/221,409
;  PRIOR FILING DATE: 2000-07-28
;  PRIOR APPLICATION NUMBER: 60/222,840
;  PRIOR FILING DATE: 2000-08-04
;  PRIOR APPLICATION NUMBER: 60/223,752
;  PRIOR FILING DATE: 2000-08-08
;  PRIOR APPLICATION NUMBER: 60/223,762
;  PRIOR FILING DATE: 2000-08-08
;  PRIOR APPLICATION NUMBER: 60/223,770
;  PRIOR FILING DATE: 2000-08-08
;  PRIOR APPLICATION NUMBER: 60/223,769
;  PRIOR FILING DATE: 2000-08-08
;  PRIOR APPLICATION NUMBER: 60/225,146
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/225,392
;  PRIOR FILING DATE: 2000-08-15
;  PRIOR APPLICATION NUMBER: 60/225,470
;  PRIOR FILING DATE: 2000-08-15
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 61
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 2752
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-624-932-1

  Query Match             92.9%;  Score 2676.4;  DB 17;  Length 2752;
  Best Local Similarity   98.9%;  Pred. No. 0;
  Matches 2728;  Conservative    0;  Mismatches   21;  Indels    9;  Gaps    3;

Qy         42 CCGCGGGGCCCCGCGCCCGGCCCGCCCGCCTGCCCGCCCGCGGCCATGGCCGTCCGGCCC 101
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCGCGGGGCCCCGCGCCCGGCCCGCCCGCCTGCCCGCCCGCGGCCATGGCCGTCCGGCCC 60

Qy        102 GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC 161
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC 120

Qy        162 CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC 221
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC 180

Qy        222 TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC 281
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC 240

Qy        282 AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG 341
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG 300

Qy        342 GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCGACCATGGAGGTCCGC 401
              ||||||||||||||||||||||||||||||||||||||   |||| |||||||||||||||
Db        301 GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCCACCATGGAGGTCCGC 360

Qy        402 ATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAG 461
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAG 420

Qy        462 TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC 521
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC 480

Qy        522 AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC 581
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC 540

Qy        582 ATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG 641
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG 600

Qy        642 AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC 701
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC 660

Qy        702 CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC 761
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC 720

Qy        762 ATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTACGTGAACGGTGGGTGG 821
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTACGTGAACGGTGGGTGG 780

Qy        822 TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG 881
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG 840

Qy        882 AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT 941
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT 900

Qy        942 GTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGCAGCTGGAGCCCGTGG 1001
              |||||  |    || | || | |    ||   | | || |||||||||||||||||||||
Db        901 GTCCAGAA---AACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGCTGGAGCCCGTGG 957

Qy       1002 AGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC 1061
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        958 AGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC 1017

Qy       1062 CCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT 1121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1018 CCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT 1077

Qy       1122 ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC 1181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1078 ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC 1137

Qy       1182 CTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCATCCTCGTTTATTGC 1241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1138 CTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCATCCTCGTTTATTGC 1197

Qy       1242 CGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC 1301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1198 CGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC 1257

Qy       1302 CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG 1361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1258 CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG 1317
```

```
Qy       1362 GACCTCAG---CACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC 1418
              ||||||||   |||||||||||||||||||||||||||||||||||||||||||||||||
Db       1318 GACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC 1377

Qy       1419 AGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC 1478
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1378 AGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC 1437

Qy       1479 ACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC 1538
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1438 ACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC 1497

Qy       1539 CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC 1598
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1498 CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC 1557

Qy       1599 TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCCAGAT 1658
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1558 TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCCAGAT 1617

Qy       1659 GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG 1718
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1618 GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG 1677

Qy       1719 AGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT 1778
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1678 AGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT 1737

Qy       1779 GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT 1838
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1738 GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT 1797

Qy       1839 GACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGCAGGATGTG 1898
              |||||||||||||||||||||||||||||||||||||||||||||||||   ||||||||
Db       1798 GACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGG---AGGATGTG 1854

Qy       1899 CTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC 1958
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1855 CTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC 1914

Qy       1959 TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG 2018
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1915 TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG 1974

Qy       2019 GCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAG 2078
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1975 GCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAG 2034

Qy       2079 TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG 2138
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2035 TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG 2094

Qy       2139 CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC 2198
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2095 CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC 2154

Qy       2199 AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAG 2258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2155 AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAG 2214

Qy       2259 CTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC 2318
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2215 CTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC 2274

Qy       2319 TTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAG 2378
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2275 TTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAG 2334

Qy       2379 CTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACC 2438
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2335 CTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACC 2394

Qy       2439 AAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG 2498
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2395 AAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG 2454

Qy       2499 GTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTG 2558
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2455 GTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTG 2514

Qy       2559 GACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGAC 2618
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2515 GACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGAC 2574

Qy       2619 AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGG 2678
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2575 AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGG 2634

Qy       2679 GAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTG 2738
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2635 GAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTG 2694

Qy       2739 GGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGAGGCCGGCCAG 2796
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2695 GGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGAGGCCGGCCAG 2752
```

```
RESULT 4
US-10-311-623-13
; Sequence 13, Application US/10311623
; Publication No. US20040023244A1
; GENERAL INFORMATION:
;  APPLICANT: INCYTE GENOMICS, INC.; GRIFFIN, Jennifer A.
;  APPLICANT:  KALLICK, Deborah A.; TRIBOULEY, Catherine M.
;  APPLICANT:  YUE, Henry; NGUYEN, Danniel B.
;  APPLICANT:  TANG, Y. Tom; LAL, Preeti G.
;  APPLICANT:  POLICKY, Jennifer L.; AZIMZAI, Yalda
;  APPLICANT:  LU, Dyung Aina M.; GRAUL, Richard C.
;  APPLICANT:  YAO, Monique G.; BURFORD, Neil
;  APPLICANT:  HAFALIA, April J. A.; BAUGHN, Mariah R.
;  APPLICANT:  BANDMAN, Olga; ARVIZU, Chandra S.
;  APPLICANT:  YANG, Junming; XU, Yuming
;  APPLICANT:  GANDHI, Ameena R.; WARREN, Bridget A.
;  APPLICANT:  DING, Li; SANJANWALA, Madhusudan M.
;  APPLICANT:  DUGGAN, Brendan M.; LU, Yan
;  TITLE OF INVENTION: RECEPTORS
;  FILE REFERENCE: PF-0793 USN
;  CURRENT APPLICATION NUMBER: US/10/311,623
;  CURRENT FILING DATE:  2002-12-17
;  PRIOR APPLICATION NUMBER: US 01/19942
;  PRIOR FILING DATE: 2001-06-21
;  PRIOR APPLICATION NUMBER: US 60/214,027
;  PRIOR FILING DATE: 2000-06-21
;  PRIOR APPLICATION NUMBER: US 60/228,045
;  PRIOR FILING DATE: 2000-08-25
;  PRIOR APPLICATION NUMBER: US 60/255,104
;  PRIOR FILING DATE: 2000-12-12
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PERL Program
; SEQ ID NO 13
;   LENGTH: 3580
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20040023244A1 6052371CB1
US-10-311-623-13

  Query Match             83.4%;  Score 2402.6;  DB 17;  Length 3580;
  Best Local Similarity   93.4%;  Pred. No. 0;
  Matches 2615;  Conservative    0;  Mismatches    9;  Indels  177;  Gaps    3;

Qy         84 GCCATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGG 143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCCATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGG 60
```

```
Qy        144 CTCCGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAAC 203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTCCGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAAC 120

Qy        204 CCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAG 263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAG 180

Qy        264 CCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGG 323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGG 240

Qy        324 GAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAG 383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||   |
Db        241 GAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTG 300

Qy        384 CCGACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTG 443
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTG 360

Qy        444 GAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAG 503
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAG 420

Qy        504 GCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAG 563
              ||||||||||||||||||   |||||||||||||||||||||||||||||||||||||||
Db        421 GCCTACATCCGCATAGCCTATTTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAG 480

Qy        564 GTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCC 623
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCC 540

Qy        624 GAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATAC 683
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATAC 600

Qy        684 ATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTAC 743
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTAC 660

Qy        744 ACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTC 803
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTC 720

Qy        804 TACGTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGG 863
              |||
Db        721 TAC--------------------------------------------------------- 723

Qy        864 CGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCT 923

Db        724 ------------------------------------------------------------ 723

Qy        924 TTCTGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGAC 983
                                                                    ||||||
Db        724 ------------------------------------------------------GTGGAC 729

Qy        984 GGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGG 1043
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        730 GGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGG 789

Qy       1044 AGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGAC 1103
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        790 AGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGAC 849

Qy       1104 CTGGACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGAC 1163
              |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db        850 CTGGACACCCGCAACTGTACCAGTGACCTCTGTGTACACACTGCTTCTGGCCCTGAGGAC 909

Qy       1164 GTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTC 1223
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        910 GTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTC 969

Qy       1224 CTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCC 1283
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        970 CTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCC 1029

Qy       1284 ATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCAT 1343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1030 ATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCAT 1089

Qy       1344 CTGCTCACCATCCAGCCGGACCTCAG---CACCACCACCACCTACCAGGGCAGTCTCTGT 1400
              ||||||||||||||||||||||||||   |||||||||||||||||||||||||||||||
Db       1090 CTGCTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGT 1149

Qy       1401 CCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCC 1460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1150 CCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCC 1209

Qy       1461 CTGGGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTC 1520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1210 CTGGGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTC 1269

Qy       1521 GTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATG 1580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1270 GTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATG 1329

Qy       1581 ACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGC 1640
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db       1330 ACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGAATCAGC 1389

Qy       1641 CTCCTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTG 1700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1390 CTCCTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTG 1449

Qy       1701 CACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATC 1760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1450 CACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATC 1509

Qy       1761 GTTAGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCAC 1820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1510 GTTAGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCAC 1569

Qy       1821 TGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGC 1880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1570 TGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGC 1629

Qy       1881 AGCTGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGC 1940
              |||||||   ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1630 AGCTGGG---AGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGC 1686

Qy       1941 CAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTG 2000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1687 CAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTG 1746

Qy       2001 GGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTG 2060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1747 GGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTG 1806

Qy       2061 GCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCA 2120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1807 GCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCA 1866

Qy       2121 CTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGG 2180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1867 CTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGG 1926

Qy       2181 GTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGC 2240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1927 GTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGC 1986

Qy       2241 TCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGG 2300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1987 TCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGG 2046

Qy       2301 AATGGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACT 2360
```

```
Db      2047 AATGGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACT 2106

Qy      2361 AGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGC 2420
Db      2107 AGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGC 2166

Qy      2421 ATCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAA 2480
Db      2167 ATCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAA 2226

Qy      2481 GCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAG 2540
Db      2227 GCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAG 2286

Qy      2541 AAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCC 2600
Db      2287 AAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCC 2346

Qy      2601 CAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCC 2660
Db      2347 CAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCC 2406

Qy      2661 ATGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCT 2720
Db      2407 ATGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCT 2466

Qy      2721 GCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAG 2780
Db      2467 GCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAG 2526

Qy      2781 TGCTGAGGCCGGCCAGGCCCGACACCTACACTCTCACCAGCTTTGGCACCCACCAAGGAC 2840
Db      2527 TGCTGAGGCCGGCCAGGCCCGACACCTACACTCTCACCAGCTTTGGCACCCACCAAGGAC 2586

Qy      2841 AGGCAGAAGCCGGACAGGGGCCCTTCCCCACACCGGGGAGA 2881
Db      2587 AGGCAGAAGCCGGACAGGGGCCCTTCCCCACACCGGGGAGA 2627


RESULT 5
US-09-933-261-1
; Sequence 1, Application US/09933261
; Publication No. US20030040046A1
;  GENERAL INFORMATION:
;       APPLICANT: Tessier-Lavigne, Marc
;                  Leonardo, E. David
;                  Hink, Lindsay
;                  Masu, Masayuki
;                  Kazuko, Keino-Masu
;       TITLE OF INVENTION: Netrin Receptors
;       NUMBER OF SEQUENCES: 8
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
;            STREET: 268 BUSH STREET, SUITE 3200
;            CITY: SAN FRANCISCO
;            STATE: CALIFORNIA
;            COUNTRY: USA
;            ZIP: 94104
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/933,261
;            FILING DATE: 20-Aug-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 08/808,982
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: OSMAN, RICHARD A
;            REGISTRATION NUMBER: 36,627
;            REFERENCE/DOCKET NUMBER: UC96-217
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (415) 343-4341
;            TELEFAX: (415) 343-4342
;  INFORMATION FOR SEQ ID NO: 1:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 3014 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: double
;            TOPOLOGY: linear
;       MOLECULE TYPE: cDNA
;       SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-933-261-1

  Query Match             77.6%;  Score 2235.8;  DB 10;  Length 3014;
  Best Local Similarity   88.5%;  Pred. No. 0;
  Matches 2474;  Conservative    0;  Mismatches  312;  Indels   11;  Gaps    4;

Qy        87 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 146
Db         1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGTGCTCCTGGGCATAGTCCTCGCCGCCTGGCTT 60

Qy       147 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 206
Db        61 CGTGGTTCGGGTGCCCAGCAGAGTGCCACGGTGGCCAATCCAGTGCCCGGTGCCAACCCC 120

Qy       207 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 266
Db       121 GACCTGCTGCCCCACTTCCTGGTAGAGCCTGAGGACGTGTACATTGTCAAGAACAAGCCG 180

Qy       267 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 326
Db       181 GTGTTGTTGGTGTGCAAGGCTGTGCCTGCCACCCAGATCTTCTTCAAGTGCAATGGGGAA 240

Qy       327 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCG 386
Db       241 TGGGTCCGCCAGGTCGATCACGTAATTGAACGCAGCACCGACAGCAGCAGCGGATTGCCA 300

Qy       387 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 446
Db       301 ACCATGGAGGTCCGTATCAACGTATCGAGGCAGCAGGTAGAGAAAGTGTTTGGGCTGGAG 360

Qy       447 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 506
Db       361 GAATACTGGTGCCAGTGTGTGGCATGGAGCTCCTCGGGTACCACCAAAAGTCAGAAGGCC 420

Qy       507 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 566
Db       421 TACATCCGGATTGCCTATTTGCGCAAGAACTTTGAGCAGGAGCCACTGGCCAAGGAAGTG 480

Qy       567 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 626
Db       481 TCACTGGAGCAAGGCATTGTACTACCTTGTCGCCCCCCAGAAGGAATCCCCCCAGCTGAG 540

Qy       627 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 686
Db       541 GTGGAGTGGCTTCGAAATGAGGACCTCGTGGACCCCTCCCTCGATCCCAATGTGTACATC 600

Qy       687 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 746
Db       601 ACGCGGGAGCACAGCCTAGTCGTGCGTCAGGCCCGCCTGGCCGACACGGCCAACTACACC 660

Qy       747 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 806
Db       661 TGTGTGGCCAAGAACATCGTAGCCCGTCGCCGAAGCACCTCTGCAGCGGTCATTGTTTAT 720

Qy       807 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 866
Db       721 GTGAACGGTGGGTGGTCGACGTGGACTGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGT 780

Qy       867 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 926
```

```
Db        781 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCACCTCTCAACGGGGGCGCCTTC 840

Qy        927 TGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGC 986
Db        841 TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACTCTG---TGCCCAGTGGATGGG 897

Qy        987 AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC 1046
Db        898 AGCTGGAGTTCGTGGAGTAAGTGGTCAGCCTGTGGGCTTGACTGCACCCACTGGCGGAGC 957

Qy       1047 CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG 1106
Db        958 CGCGAGTGCTCTGACCCAGCACCCCGCAATGGAGGTGAGGAGTGTCGGGGTGCTGACCTG 1017

Qy       1107 GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG 1166
Db       1018 GACACCCGCAACTGTACCAGTGACCTCTGCCTGCACACCGCTTCTTGCCCCGAGGACGTG 1077

Qy       1167 GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC 1226
Db       1078 GCTCTCTACATCGGCCTTGTCGCTGTGGCTGTGTGCCTCTTCTTGCTGTTGCTGGCCCTT 1137

Qy       1227 ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1286
Db       1138 GGACTCATTTACTGTCGCAAGAAGGAAGGGCTGGACTCCGATGTGGCCGACTCGTCCATC 1197

Qy       1287 CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1346
Db       1198 CTCACCTCGGGCTTCCAGCCTGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCACCTG 1257

Qy       1347 CTCACCATCCAGCCGGACCTCAGCACCACC---ACCACCTACCAGGGCAGTCTCTGTCCC 1403
Db       1258 CTCACCATCCAGCCAGACCTCAGCACCACCACTACCACCTACCAGGGCAGTCTATGTTCG 1317

Qy       1404 CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG 1463
Db       1318 AGGCAGGATGGACCCAGCCCCAAGTTCCAGCTCTCTAATGGTCACCTGCTCAGCCCACTG 1377

Qy       1464 GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC 1523
Db       1378 GGGAGTGGCCGCCATACGTTGCACCACAGCTCACCCACCTCTGAGGCTGAGGACTTCGTC 1437

Qy       1524 TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC 1583
Db       1438 TCCCGCCTCTCCACCCAAAACTACTTTCGTTCCCTGCCCCGCGGCACCAGCAACATGGCC 1497

Qy       1584 TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC 1643
Db       1498 TACGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACGGGGATCAGCCTC 1557

Qy       1644 CTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC 1703
Db       1558 CTCATACCCCCGGATGCCATCCCCCGAGGAAAGATCTACGAGATCTACCTCACACTGCAC 1617

Qy       1704 AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT 1763
Db       1618 AAGCCAGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCAGTCGTT 1677

Qy       1764 AGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT 1823
Db       1678 AGCTGTGGGCCCCCAGGAGTCCTGCTCACCCGGCCAGTCATCCTTGCAATGGACCACTGT 1737

Qy       1824 GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC 1883
Db       1738 GGAGAGCCCAGCCCTGACAGCTGGAGTCTGCGCCTCAAAAAGCAGTCCTGCGAGGGCAGT 1797

Qy       1884 TGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG 1943
Db       1798 TGGG---AGGATGTGCTGCACCTTGGTGAGGAGTCACCTTCCCACCTCTACTACTGCCAG 1854

Qy       1944 CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA 2003
Db       1855 CTGGAGGCCGGGGCCTGCTATGTCTTCACGGAGCAGCTGGGCCGCTTTGCCCTGGTAGGA 1914

Qy       2004 GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC 2063
Db       1915 GAGGCCCTCAGCGTGGCTGCCACCAAGCGCCTCAGGCTCCTTCTGTTTGCTCCCGTGGCC 1974

Qy       2064 TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC 2123
Db       1975 TGTACGTCCCTTGAGTACAACATCCGAGTGTACTGCCTACACGACACCCACGACGCTCTC 2034

Qy       2124 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTC 2183
Db       2035 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTAGGTGGACAGCTGATCCAGGAGCCTCGCGTC 2094

Qy       2184 CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCC 2243
Db       2095 CTGCACTTCAAAGACAGTTACCACAACCTACGTCTCTCCATCCACGACGTGCCCAGCTCC 2154

Qy       2244 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2303
Db       2155 CTGTGGAAGAGCAAGCTACTTGTCAGCTACCAGGAGATCCCTTTTTACCACATCTGGAAC 2214

Qy       2304 GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2363
Db       2215 GGCACCCAGCAGTATCTGCACTGCACCTTCACCCTGGAGCGCATCAACGCCAGCACCAGC 2274

Qy       2364 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2423
Db       2275 GACCTGGCCTGCAAGGTGTGGGTGTGGCAGGTGGAGGGAGATGGGCAGAGCTTCAACATC 2334

Qy       2424 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2483
Db       2335 AACTTCAACATCACTAAGGACACAAGGTTTGCTGAATTGTTGGCTCTGGAGAGTGAAGGG 2394

Qy       2484 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2543
Db       2395 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAAAAG 2454

Qy       2544 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2603
Db       2455 ATCATCGCCAGTCTGGACCCACCCTGCAGCCGGGGCGCCGACTGGAGAACTCTAGCCCAG 2514

Qy       2604 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG 2663
Db       2515 AAACTTCACCTGGACAGCCATCTTAGCTTCTTTGCCTCCAAGCCCAGCCCTACAGCCATG 2574

Qy       2664 ATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2723
Db       2575 ATCCTCAACCTATGGAGGCACGGCACTTCCCCAACGGCAACCTCGGCCAGCTGGCAGCA 2634

Qy       2724 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2783
Db       2635 GCTGTGGCCGGACTGGGCCAACCAGATGCTGGCCTCTTCACGGTGTCGGAGGCCGAGTGT 2694

Qy       2784 TGAGGCCGGCCAGGCCCG--ACACCTACACTCTCACCAGCTTTGGCACCCACCAAGGACA 2841
Db       2695 TGAGACCAGCCAGGCCGGTAATGCCTACATTCTCACCAGCTTTGACACCTGCCAGGGACA 2754

Qy       2842 GGCAGAAGCCGGACAGGGGCCCTTCCCCACACCGGGG 2878
Db       2755 GGCAAAACCAGACAGGGGCCCTTCCCCCACACCCGGG 2791
```

RESULT 6
US-10-256-702-1
; Sequence 1, Application US/10256702
; Publication No. US20030059859A1
;   GENERAL INFORMATION:
;        APPLICANT: Tessier-Lavigne, Marc
;                   Leonardo, E. David
;                   Hink, Lindsay
;                   Masu, Masayuki
;                   Kazuko, Keino-Masu
;        TITLE OF INVENTION: Netrin Receptors

```
;        NUMBER OF SEQUENCES: 8
;        CORRESPONDENCE ADDRESS:
;            ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
;            STREET: 268 BUSH STREET, SUITE 3200
;            CITY: SAN FRANCISCO
;            STATE: CALIFORNIA
;            COUNTRY: USA
;            ZIP: 94104
;        COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/256,702
;            FILING DATE: 27-Sep-2002
;            CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/09/933,261
;            FILING DATE: 20-Aug-2001
;            APPLICATION NUMBER: 08/808,982
;            FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;            NAME: OSMAN, RICHARD A
;            REGISTRATION NUMBER: 36,627
;            REFERENCE/DOCKET NUMBER: UC96-217
;        TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (415) 343-4341
;            TELEFAX: (415) 343-4342
;    INFORMATION FOR SEQ ID NO: 1:
;        SEQUENCE CHARACTERISTICS:
;            LENGTH: 3014 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: double
;            TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-256-702-1

  Query Match             77.6%;  Score 2235.8;  DB 15;  Length 3014;
  Best Local Similarity   88.5%;  Pred. No. 0;
  Matches 2474;  Conservative    0;  Mismatches  312;  Indels   11;  Gaps    4;

Qy         87 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 146
Db          1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGTGCTCCTGGGCATAGTCCTCGCCGCCTGGCTT 60

Qy        147 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 206
Db         61 CGTGGTTCGGGTGCCCAGCAGAGTGCCACGGTGGCCAATCCAGTGCCCGGTGCCAACCCC 120

Qy        207 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 266
Db        121 GACCTGCTGCCCCACTTCCTGGTAGAGCCTGAGGACGTGTACATTGTCAAGAACAAGCCG 180

Qy        267 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 326
Db        181 GTGTTGTTGGTGTGCAAGGCTGTGCCTGCCACCCAGATCTTCTTCAAGTGCAATGGGGAA 240

Qy        327 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCG 386
Db        241 TGGGTCCGCCAGGTCGATCACGTAATTGAACGCAGCACCGACAGCAGCAGCGGATTGCCA 300

Qy        387 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 446
Db        301 ACCATGGAGGTCCGTATCAACGTATCGAGGCAGCAGGTAGAGAAAGTGTTTGGGCTGGAG 360

Qy        447 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 506
Db        361 GAATACTGGTGCCAGTGTGTGGCATGGAGCTCCTCGGGTACCACCAAAAGTCAGAAGGCC 420

Qy        507 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 566
Db        421 TACATCCGGATTGCCTATTTGCGCAAGAACTTTGAGCAGGAGCCACTGGCCAAGGAAGTG 480

Qy        567 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 626
Db        481 TCACTGGAGCAAGGCATTGTACTACCTTGTCGCCCCCCAGAAGGAATCCCCCCAGCTGAG 540

Qy        627 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 686
Db        541 GTGGAGTGGCTTCGAAATGAGGACCTCGTGGACCCCTCCCTCGATCCCAATGTGTACATC 600

Qy        687 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 746
Db        601 ACGCGGGAGCACAGCCTAGTCGTGCGTCAGGCCCGCCTGGCCGACACGGCCAACTACACC 660

Qy        747 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 806
Db        661 TGTGTGGCCAAGAACATCGTAGCCCGTCGCCGAAGCACCTCTGCAGCGGTCATTGTTTAT 720

Qy        807 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 866
Db        721 GTGAACGGTGGGTGGTCGACGTGGACTGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGT 780

Qy        867 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 926
Db        781 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCACCTCTCAACGGGGGCGCCTTC 840

Qy        927 TGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGC 986
Db        841 TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACTCTG---TGCCCAGTGGATGGG 897

Qy        987 AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC 1046
Db        898 AGCTGGAGTTCGTGGAGTAAGTGGTCAGCCTGTGGGCTTGACTGCACCCACTGGCGGAGC 957

Qy       1047 CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG 1106
Db        958 CGCGAGTGCTCTGACCCAGCACCCCGCAATGGAGGTGAGGAGTGTCGGGGTGCTGACCTG 1017

Qy       1107 GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG 1166
Db       1018 GACACCCGCAACTGTACCAGTGACCTCTGCCTGCACACCGCTTCTTGCCCCGAGGACGTG 1077

Qy       1167 GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC 1226
Db       1078 GCTCTCTACATCGGCCTTGTCGCTGTGGCTGTGTGCCTCTTCTTGCTGTTGCTGGCCCTT 1137

Qy       1227 ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1286
Db       1138 GGACTCATTTACTGTCGCAAGAAGGAAGGGCTGGACTCCGATGTGGCCGACTCGTCCATC 1197

Qy       1287 CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1346
Db       1198 CTCACCTCGGGCTTCCAGCCTGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCACCTG 1257

Qy       1347 CTCACCATCCAGCCGGACCTCAGCACCACC---ACCACCTACCAGGGCAGTCTCTGTCCC 1403
Db       1258 CTCACCATCCAGCCAGACCTCAGCACCACCACTACCACCTACCAGGGCAGTCTATGTTCG 1317

Qy       1404 CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG 1463
Db       1318 AGGCAGGATGGACCCAGCCCCAAGTTCCAGCTCTCTAATGGTCACCTGCTCAGCCCACTG 1377

Qy       1464 GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC 1523
Db       1378 GGAGTGGCCGCCATACGTTGCACCACAGCTCACCCACCTCTGAGGCTGAGGACTTCGTC 1437

Qy       1524 TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC 1583
Db       1438 TCCCGCCTCTCCACCCAAAACTACTTTCGTTCCCTGCCCCGCGGCACCAGCAACATGGCC 1497

Qy       1584 TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC 1643
Db       1498 TACGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACGGGGATCAGCCTC 1557
```

```
Qy       1644 CTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC 1703
Db       1558 CTCATACCCCCGGATGCCATCCCCCGAGGAAAGATCTACGAGATCTACCTCACACTGCAC 1617

Qy       1704 AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT 1763
Db       1618 AAGCCAGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCAGTCGTT 1677

Qy       1764 AGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT 1823
Db       1678 AGCTGTGGGCCCCCAGGAGTCCTGCTCACCCGGCCAGTCATCCTTGCAATGGACCACTGT 1737

Qy       1824 GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC 1883
Db       1738 GGAGAGCCCAGCCCTGACAGCTGGAGTCTGCGCCTCAAAAAGCAGTCCTGCGAGGGCAGT 1797

Qy       1884 TGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG 1943
Db       1798 TGGG---AGGATGTGCTGCACCTTGGTGAGGAGTCACCTTCCCACCTCTACTACTGCCAG 1854

Qy       1944 CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA 2003
Db       1855 CTGGAGGCCGGGGCCTGCTATGTCTTCACGGAGCAGCTGGGCCGCTTTGCCCTGGTAGGA 1914

Qy       2004 GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC 2063
Db       1915 GAGGCCCTCAGCGTGGCTGCCACCAAGCGCCTCAGGCTCCTTCTGTTTGCTCCCGTGGCC 1974

Qy       2064 TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC 2123
Db       1975 TGTACGTCCCTTGAGTACAACATCCGAGTGTACTGCCTACACGACACCCACGACGCTCTC 2034

Qy       2124 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTC 2183
Db       2035 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTAGGTGGACAGCTGATCCAGGAGCCTCGCGTC 2094

Qy       2184 CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCC 2243
Db       2095 CTGCACTTCAAAGACAGTTACCACAACCTACGTCTCTCCATCCACGACGTGCCCAGCTCC 2154

Qy       2244 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2303
Db       2155 CTGTGGAAGAGCAAGCTACTTGTCAGCTACCAGGAGATCCCTTTTTACCACATCTGGAAC 2214

Qy       2304 GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2363
Db       2215 GGCACCCAGCAGTATCTGCACTGCACCTTCACCCTGGAGCGCATCAACGCCAGCACCAGC 2274

Qy       2364 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2423
Db       2275 GACCTGGCCTGCAAGGTGTGGGTGTGGCAGGTGGAGGGAGATGGGCAGAGCTTCAACATC 2334

Qy       2424 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2483
Db       2335 AACTTCAACATCACTAAGGACACAAGGTTTGCTGAATTGTTGGCTCTGGAGAGTGAAGGG 2394

Qy       2484 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2543
Db       2395 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAAAAG 2454

Qy       2544 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2603
Db       2455 ATCATCGCCAGTCTGGACCCACCCTGCAGCCGGGGCGCCGACTGGAGAACTCTAGCCCAG 2514

Qy       2604 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG 2663
Db       2515 AAACTTCACCTGGACAGCCATCTTAGCTTCTTTGCCTCCAAGCCCAGCCCTACAGCCATG 2574

Qy       2664 ATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2723
Db       2575 ATCCTCAACCTATGGGAGGCACGGCACTTCCCCAACGGCAACCTCGGCCAGCTGGCAGCA 2634

Qy       2724 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2783
Db       2635 GCTGTGGCCGGACTGGGCCAACCAGATGCTGGCCTCTTCACGGTGTCGGAGGCCGAGTGT 2694

Qy       2784 TGAGGCCGGCCAGGCCCG--ACACCTACACTCTCACCAGCTTTGGCACCCACCAAGGACA 2841
Db       2695 TGAGACCAGCCAGGCCGGTAATGCCTACATTCTCACCAGCTTTGACACCTGCCAGGGACA 2754

Qy       2842 GGCAGAAGCCGGACAGGGGCCCTTCCCCACACCGGGG 2878
Db       2755 GGCAAAACCAGACAGGGGCCCTTCCCCCACACCCGGG 2791
```

RESULT 7
US-10-240-154-15
; Sequence 15, Application US/10240154
; Publication No. US20030175741A1
; GENERAL INFORMATION:
;  APPLICANT: Cochran et al.
;  TITLE OF INVENTION: SCHIZOPHRENIA RELATED GENES
;  FILE REFERENCE: CKFW-P01-006
;  CURRENT APPLICATION NUMBER: US/10/240,154
;  CURRENT FILING DATE:  2001-04-02
;  PRIOR APPLICATION NUMBER: PCT/GB01/01486
;  PRIOR FILING DATE: 2001-04-02
;  NUMBER OF SEQ ID NOS: 34
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 15
;   LENGTH: 2697
;   TYPE: DNA
;   ORGANISM: Rattus sp.
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(2697)
US-10-240-154-15

```
  Query Match             76.0%;  Score 2189.4;  DB 15;  Length 2697;
  Best Local Similarity   88.9%;  Pred. No. 0;
  Matches 2403;  Conservative    0;  Mismatches  291;  Indels    9;  Gaps    3;

Qy         87 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 146
Db          1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGTGCTCCTGGGCATAGTCCTCGCCGCCTGGCTT 60

Qy        147 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 206
Db         61 CGTGGTTCGGGTGCCCAGCAGAGTGCCACGGTGGCCAATCCAGTGCCCGGTGCCAACCCC 120

Qy        207 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 266
Db        121 GACCTGCTGCCCCACTTCCTGGTAGAGCCTGAGGACGTGTACATTGTCAAGAACAAGCCG 180

Qy        267 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 326
Db        181 GTGTTGTTGGTGTGCAAGGCTGTGCCTGCCACCCAGATCTTCTTCAAGTGCAATGGGGAA 240

Qy        327 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCG 386
Db        241 TGGGTCCGCCAGGTCGATCACGTAATTGAACGCAGCACCGACAGCAGCAGCGGATTGCCA 300

Qy        387 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 446
Db        301 ACCATGGAGGTCCGTATCAACGTATCGAGGCAGCAGGTAGAGAAAGTGTTTGGGCTGGAG 360

Qy        447 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 506
Db        361 GAATACTGGTGCCAGTGTGTGGCATGGAGCTCCTCGGGTACCACCAAAAGTCAGAAGGCC 420

Qy        507 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 566
Db        421 TACATCCGGATTGCCTATTTGCGCAAGAACTTTGAGCAGGAGCCACTGGCCAAGGAAGTG 480

Qy        567 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 626
```

```
Db     481 TCACTGGAGCAAGGCATTGTACTACCTTGTCGCCCCCCAGAAGGAATCCCCCCAGCTGAG 540

Qy     627 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 686
Db     541 GTGGAGTGGCTTCGAAATGAGGACCTCGTGGACCCCTCCCTCGATCCCAATGTGTACATC 600

Qy     687 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 746
Db     601 ACGCGGGAGCACAGCCTAGTCGTGCGTCAGGCCCGCCTGGCCGACACGGCCAACTACACC 660

Qy     747 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 806
Db     661 TGTGTGGCCAAGAACATCGTAGCCCGTCGCCGAAGCACCTCTGCAGCGGTCATTGTTTAT 720

Qy     807 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 866
Db     721 GTGAACGGTGGGTGGTCGACGTGGACTGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGT 780

Qy     867 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 926
Db     781 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCACCTCTCAACGGGGGCGCCTTC 840

Qy     927 TGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGC 986
Db     841 TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACTCTG---TGCCCAGTGGATGGG 897

Qy     987 AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC 1046
Db     898 AGCTGGAGTTCGTGGAGTAAGTGGTCAGCCTGTGGGCTTGACTGCACCCACTGGCGGAGC 957

Qy    1047 CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG 1106
Db     958 CGCGAGTGCTCTGACCCAGCACCCCGCAATGGAGGTGAGGAGTGTCGGGGTGCTGACCTG 1017

Qy    1107 GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG 1166
Db    1018 GACACCCGCAACTGTACCAGTGACCTCTGCCTGCACACCGCTTCTTGCCCCGAGGACGTG 1077

Qy    1167 GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC 1226
Db    1078 GCTCTCTACATCGGCCTTGTCGCTGTGGCTGTGTGCCTCTTCTTGCTGTTGCTGGCCCTT 1137

Qy    1227 ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1286
Db    1138 GGACTCATTTACTGTCGCAAGAAGGAAGGGCTGGACTCCGATGTGGCCGACTCGTCCATC 1197

Qy    1287 CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1346
Db    1198 CTCACCTCGGGCTTCCAGCCTGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCACCTG 1257

Qy    1347 CTCACCATCCAGCCGGACCTCAGCACCACC---ACCACCTACCAGGGCAGTCTCTGTCCC 1403
Db    1258 CTCACCATCCAGCCAGACCTCAGCACCACCACTACCACCTACCAGGGCAGTCTATGTTCG 1317

Qy    1404 CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG 1463
Db    1318 AGGCAGGATGGACCCAGCCCCAAGTTCCAGCTCTCTAATGGTCACCTGCTCAGCCCACTG 1377

Qy    1464 GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC 1523
Db    1378 GGGAGTGGCCGCCATACGTTGCACCACAGCTCACCCACCTCTGAGGCTGAGGACTTCGTC 1437

Qy    1524 TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC 1583
Db    1438 TCCCGCCTCTCCACCCAAAACTACTTTCGTTCCCTGCCCCGCGGCACCAGCAACATGGCC 1497

Qy    1584 TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC 1643
Db    1498 TACGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACGGGGATCAGCCTC 1557

Qy    1644 CTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC 1703
Db    1558 CTCATACCCCCGGATGCCATCCCCCGAGGAAAGATCTACGAGATCTACCTCACACTGCAC 1617

Qy    1704 AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT 1763
Db    1618 AAGCCAGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCAGTCGTT 1677

Qy    1764 AGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT 1823
Db    1678 AGCTGTGGGCCCCCAGGAGTCCTGCTCACCCGGCCAGTCATCCTTGCAATGGACCACTGT 1737

Qy    1824 GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC 1883
Db    1738 GGAGAGCCCAGCCCTGACAGCTGGAGTCTGCGCCTCAAAAAGCAGTCCTGCGAGGGCAGT 1797

Qy    1884 TGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG 1943
Db    1798 TGGG---AGGATGTGCTGCACCTTGGTGAGGAGTCACCTTCCCACCTCTACTACTGCCAG 1854

Qy    1944 CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA 2003
Db    1855 CTGGAGGCCGGGGCCTGCTATGTCTTCACGGAGCAGCTGGGCCGCTTTGCCCTGGTAGGA 1914

Qy    2004 GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC 2063
Db    1915 GAGGCCCTCAGCGTGGCTGCCACCAAGCGCCTCAGGCTCCTTCTGTTTGCTCCCGTGGCC 1974

Qy    2064 TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC 2123
Db    1975 TGTACGTCCCTTGAGTACAACATCCGAGTGTACTGCCTACACGACACCCACGACGCTCTC 2034

Qy    2124 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTC 2183
Db    2035 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTAGGTGGACAGCTGATCCAGGAGCCTCGCGTC 2094

Qy    2184 CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCC 2243
Db    2095 CTGCACTTCAAAGACAGTTACCACAACCTACGTCTCTCCATCCACGACGTGCCCAGCTCC 2154

Qy    2244 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2303
Db    2155 CTGTGGAAGAGCAAGCTACTTGTCAGCTACCAGGAGATCCCTTTTTACCACATCTGGAAC 2214

Qy    2304 GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2363
Db    2215 GGCACCCAGCAGTATCTGCACTGCACCTTCACCCTGGAGCGCATCAACGCCAGCACCAGC 2274

Qy    2364 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2423
Db    2275 GACCTGGCCTGCAAGGTGTGGGTGTGGCAGGTGGAGGGAGATGGGCAGAGCTTCAACATC 2334

Qy    2424 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2483
Db    2335 AACTTCAACATCACTAAGGACACAAGGTTTGCTGAATTGTTGGCTCTGGAGAGTGAAGGG 2394

Qy    2484 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2543
Db    2395 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAAAAG 2454

Qy    2544 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2603
Db    2455 ATCATCGCCAGTCTGGACCCACCCTGCAGCCGGGGCGCCGACTGGAGAACTCTAGCCCAG 2514

Qy    2604 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG 2663
Db    2515 AAACTTCACCTGGACAGCCATCTTAGCTTCTTTGCCTCCAAGCCCAGCCCTACAGCCATG 2574

Qy    2664 ATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2723
Db    2575 ATCCTCAACCTATGGAGGCACGGCACTTCCCCAACGGCAACCTCGGCCAGCTGGCAGCA 2634

Qy    2724 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2783
Db    2635 GCTGTGGCCGGACTGGGCCAACCAGATGCTGGCCTCTTCACGGTGTCGGAGGCCGAGTGT 2694
```

```
Qy       2784 TGA 2786
              |||
Db       2695 TGA 2697

RESULT 8
US-09-933-261-2
; Sequence 2, Application US/09933261
; Publication No. US20030040046A1
;  GENERAL INFORMATION:
;       APPLICANT: Tessier-Lavigne, Marc
;                  Leonardo, E. David
;                  Hink, Lindsay
;                  Masu, Masayuki
;                  Kazuko, Keino-Masu
;       TITLE OF INVENTION: Netrin Receptors
;       NUMBER OF SEQUENCES: 8
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
;            STREET: 268 BUSH STREET, SUITE 3200
;            CITY: SAN FRANCISCO
;            STATE: CALIFORNIA
;            COUNTRY: USA
;            ZIP: 94104
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/933,261
;            FILING DATE: 20-Aug-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 08/808,982
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: OSMAN, RICHARD A
;            REGISTRATION NUMBER: 36,627
;            REFERENCE/DOCKET NUMBER: UC96-217
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (415) 343-4341
;            TELEFAX: (415) 343-4342
;  INFORMATION FOR SEQ ID NO: 2:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 1787 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: double
;            TOPOLOGY: linear
;       MOLECULE TYPE: cDNA
;       SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-933-261-2

  Query Match             54.5%;  Score 1570.4;  DB 10;  Length 1787;
  Best Local Similarity   96.9%;  Pred. No. 0;
  Matches 1721;  Conservative    0;  Mismatches   36;  Indels   19;  Gaps   11;

Qy       1114 GCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCT 1173
              |||||||||||||||||||||| |||||| |||||||||||||||||||||||||||||
Db          1 GCAACTGTACCAGTGACCTCTG-GTACACACTGCTTCTGGCCCTGAGGACGTGGCCCTCT 59

Qy       1174 ATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCATCCTCG 1233
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         60 ATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCATCCTCG 119

Qy       1234 TTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCT 1293
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        120 TTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCT 179

Qy       1294 CAGGCTTCCAGCCCGTCAGCATC-AAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACC 1352
              ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        180 CAGGCTTCCAGCCCGTCAGCATCTAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACC 239

Qy       1353 ATCCAGCCGGACCTCAG---CACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAG 1409
              |||||||||||||||||   ||||||||||||||||||||||||||||||||||||||||
Db        240 ATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAG 299

Qy       1410 GATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGC 1469
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        300 GATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGC 359

Qy       1470 GGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGC 1529
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        360 GGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGC 419

Qy       1530 CTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGG 1589
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        420 CTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGG 479

Qy       1590 ACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATC 1649
              |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db        480 ACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGAATCAGCCTCCTCATC 539

Qy       1650 CCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCG 1709
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 CCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCG 599

Qy       1710 GAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGT 1769
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 GAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGT 659

Qy       1770 GGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAG 1829
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 GGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAG 719

Qy       1830 CCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAG 1889
              ||||||||||||||||||||||| ||||||||||||||||||||||||||| ||||||||
Db        720 CCCAGCCCTGACAGCTGGAGCCTGGCCCTCAAAAAGCAGTCGTGCGAGGG-AGCTGGGAG 778

Qy       1890 CAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG 1949
              | |     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        779 GATG----TCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG 834

Qy       1950 GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC 2009
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        835 GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC 894

Qy       2010 CTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC 2069
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        895 CTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC 954

Qy       2070 TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG 2129
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        955 TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG 1014

Qy       2130 GTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC 2189
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1015 GTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC 1074

Qy       2190 TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG 2249
              || ||||||||||||||||||||  ||| ||||||||||||||||||||||||||||||
Db       1075 TT-AAGGACAGTTACCACAACCT--GCCCTATCATCCACGATGTGCCCAGCTCCCTGTGG 1131

Qy       2250 AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG 2309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1132 AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG 1191

Qy       2310 CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG 2369
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1192 CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG 1251

Qy       2370 GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC 2429
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1252 GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC 1311
```

```
Qy      2430 AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC 2489
Db      1312 AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC 1371

Qy      2490 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT 2549
Db      1372 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT 1431

Qy      2550 TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC 2609
Db      1432 TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC 1491

Qy      2610 CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC 2669
Db      1492 CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC 1551

Qy      2670 AACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG 2729
Db      1552 AACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG 1611

Qy      2730 GCTGGACTGGGCCAGCCAGACGCTGGCCTC-TTCACAGTG-TCGGAGGCTGAGTGCTGAG 2787
Db      1612 GCTGGGACTGGCCAGCAGGACGGTGGCTTCTTTCACAGTGTTCGGAGGCTGAGTGCTGAG 1671

Qy      2788 GCCGGCCAGGCCCGACACCTACACTCTCACCAGCTTTGGCA--CCCACCAAGGACAGGCA 2845
Db      1672 GCCGGCCAGG--CGAACACTACAATTTTACCAGTTTTGGGAACCCACCAAGGGACAGGCA 1729

Qy      2846 GAAGCCGGACAGGGGCCCTTCCCCACACCGGGGAGA 2881
Db      1730 GAAGCCGGACAAGGGCTTTTCCCAAAACCGGGGAGA 1765

RESULT 9
US-10-256-702-2
; Sequence 2, Application US/10256702
; Publication No. US20030059859A1
;  GENERAL INFORMATION:
;       APPLICANT: Tessier-Lavigne, Marc
;                  Leonardo, E. David
;                  Hink, Lindsay
;                  Masu, Masayuki
;                  Kazuko, Keino-Masu
;       TITLE OF INVENTION: Netrin Receptors
;       NUMBER OF SEQUENCES: 8
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
;            STREET: 268 BUSH STREET, SUITE 3200
;            CITY: SAN FRANCISCO
;            STATE: CALIFORNIA
;            COUNTRY: USA
;            ZIP: 94104
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/256,702
;            FILING DATE: 27-Sep-2002
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/09/933,261
;            FILING DATE: 20-Aug-2001
;            APPLICATION NUMBER: 08/808,982
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: OSMAN, RICHARD A
;            REGISTRATION NUMBER: 36,627
;            REFERENCE/DOCKET NUMBER: UC96-217
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (415) 343-4341
;            TELEFAX: (415) 343-4342
;
;   INFORMATION FOR SEQ ID NO: 2:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 1787 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: double
;            TOPOLOGY: linear
;       MOLECULE TYPE: cDNA
;       SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-256-702-2

  Query Match             54.5%;  Score 1570.4;  DB 15;  Length 1787;
  Best Local Similarity   96.9%;  Pred. No. 0;
  Matches 1721;  Conservative    0;  Mismatches   36;  Indels   19;  Gaps   11;

Qy      1114 GCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCT 1173
Db         1 GCAACTGTACCAGTGACCTCTG-GTACACACTGCTTCTGGCCCTGAGGACGTGGCCCTCT 59

Qy      1174 ATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCATCCTCG 1233
Db        60 ATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCATCCTCG 119

Qy      1234 TTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCT 1293
Db       120 TTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCT 179

Qy      1294 CAGGCTTCCAGCCCGTCAGCATC-AAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACC 1352
Db       180 CAGGCTTCCAGCCCGTCAGCATCTAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACC 239

Qy      1353 ATCCAGCCGGACCTCAG---CACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAG 1409
Db       240 ATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAG 299

Qy      1410 GATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGC 1469
Db       300 GATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGC 359

Qy      1470 GGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGC 1529
Db       360 GGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGC 419

Qy      1530 CTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGG 1589
Db       420 CTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGG 479

Qy      1590 ACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATC 1649
Db       480 ACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGAATCAGCCTCCTCATC 539

Qy      1650 CCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCG 1709
Db       540 CCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCG 599

Qy      1710 GAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGT 1769
Db       600 GAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGT 659

Qy      1770 GGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAG 1829
Db       660 GGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAG 719

Qy      1830 CCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAG 1889
Db       720 CCCAGCCCTGACAGCTGGAGCCTGGCCCTCAAAAAGCAGTCGTGCGAGGG-AGCTGGGAG 778

Qy      1890 CAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG 1949
Db       779 GATG----TCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG 834

Qy      1950 GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC 2009
Db       835 GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC 894
```

```
Qy     2010 CTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC 2069
Db      895 CTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC 954

Qy     2070 TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG 2129
Db      955 TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG 1014

Qy     2130 GTGGTGCAGCTGGAGAAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC 2189
Db     1015 GTGGTGCAGCTGGAGAAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC 1074

Qy     2190 TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG 2249
Db     1075 TT-AAGGACAGTTACCACAACCT--GCCCTATCATCCACGATGTGCCCAGCTCCCTGTGG 1131

Qy     2250 AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG 2309
Db     1132 AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG 1191

Qy     2310 CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG 2369
Db     1192 CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG 1251

Qy     2370 GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC 2429
Db     1252 GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC 1311

Qy     2430 AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC 2489
Db     1312 AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC 1371

Qy     2490 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT 2549
Db     1372 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT 1431

Qy     2550 TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC 2609
Db     1432 TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC 1491

Qy     2610 CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC 2669
Db     1492 CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC 1551

Qy     2670 AACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG 2729
Db     1552 AACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG 1611

Qy     2730 GCTGGACTGGGCCAGCCAGACGCTGGCCTC-TTCACAGTG-TCGGAGGCTGAGTGCTGAG 2787
Db     1612 GCTGGGACTGGCCAGCAGGACGGTGGCTTCTTTCACAGTGTTCGGAGGCTGAGTGCTGAG 1671

Qy     2788 GCCGGCCAGGCCCGACACCTACACTCTCACCAGCTTTGGCA--CCCACCAAGGACAGGCA 2845
Db     1672 GCCGGCCAGG--CGAACACTACAATTTTACCAGTTTTGGGAACCCACCAAGGACAGGCA 1729

Qy     2846 GAAGCCGGACAGGGGCCCTTCCCCACACCGGGGAGA 2881
Db     1730 GAAGCCGGACAAGGGCTTTTCCCAAAACCGGGGAGA 1765
```

RESULT 10
US-10-296-115-365
; Sequence 365, Application US/10296115
; Publication No. US20040053248A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq Inc
;  TITLE OF INVENTION: No. US20040053248A1el Nucleic Acids and Polypeptides
;  FILE REFERENCE: 784PCT
;  CURRENT APPLICATION NUMBER: US/10/296,115
;  CURRENT FILING DATE:  2002-11-18
;  PRIOR APPLICATION NUMBER: US09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  PRIOR APPLICATION NUMBER: US09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  NUMBER OF SEQ ID NOS: 1478
; SEQ ID NO 365
;   LENGTH: 1321
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-296-115-365

```
  Query Match            41.4%;  Score 1193.6;  DB 13;  Length 1321;
  Best Local Similarity  97.8%;  Pred. No. 2.6e-287;
  Matches 1295;  Conservative    0;  Mismatches   19;  Indels   10;  Gaps    8;

Qy     1476 CACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCC 1535
Db        1 CACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCC 60

Qy     1536 ACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTC 1595
Db       61 ACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTC 120

Qy     1596 AACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCCA 1655
Db      121 AACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGAATCAGCCTCCTCATCCCCCCA 180

Qy     1656 GATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGAC 1715
Db      181 GATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGAC 240

Qy     1716 GTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCC 1775
Db      241 GTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCC 300

Qy     1776 CCT-GGCGTCCTGCTCACCCGGCCAGTCATCCT-GGCTATGGACCACTGT-GGGGAGCCC 1832
Db      301 CCTGGGCGTCCTGCTTACCCGGCCAGTCATCCTGGGGTATGGACCACTGTGGGGGAGCCC 360

Qy     1833 AGCCCTGACAGCT-GGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGCA 1891
Db      361 AGCCCTGACAGCTGGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGG---A 417

Qy     1892 GGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGC 1951
Db      418 GGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGC 477

Qy     1952 CAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCT 2011
Db      478 CAGTGCCTGCTACGTCTTCACCGAGCAGCTGAGCCGCTATGCCCTGGTGGGAGAGGCCCT 537

Qy     2012 CAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTC 2071
Db      538 CAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTC 597

Qy     2072 CCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGT 2131
Db      598 CCTCGAGTACAACATACTGGTCTACTGCCTGCATGACACTCACGATGCACTCAACGTAGT 657

Qy     2132 GGTGCAGCTGGAGAAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTT 2191
Db      658 GGTGCAGCTGGAGAAGCAGCTGCAGGGACAGCTGATCCAGGAGCCACTGGTACTGCACTT 717

Qy     2192 CAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAA 2251
Db      718 CAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAA 777

Qy     2252 GAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCA 2311
Db      778 GAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCA 837

Qy     2312 GCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGC 2371
Db      838 GCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGC 897
```

```
Qy       2372 CTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAA 2431
Db        898 CTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAA 957

Qy       2432 CATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCC 2491
Db        958 CATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCC 1017

Qy       2492 AGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTC 2551
Db       1018 AGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTC 1077

Qy       2552 CAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCA 2611
Db       1078 CAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCA 1137

Qy       2612 CCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAA 2671
Db       1138 CCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAA 1197

Qy       2672 CCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGC 2731
Db       1198 CCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGC 1257

Qy       2732 T-GGACTGGGCCAGCCAGACGCTGGCCTC-TTCACAGTG-TCGGAGGCTGAGTGCTGAGG 2788
Db       1258 TGGGACTGGGCCAGCAGGACGGTGGCTTCTTTCACAGTGTTCGGAGGCTGAGTGCTGAGG 1317

Qy       2789 CCGG 2792
Db       1318 CCGG 1321
```

RESULT 11
US-09-833-381-1806
; Sequence 1806, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
;  APPLICANT: Robison, Keith E.
;  TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
;  FILE REFERENCE: 5800-119
;  CURRENT APPLICATION NUMBER: US/09/833,381
;  CURRENT FILING DATE:  2001-04-11
;  PRIOR APPLICATION NUMBER: 09/516,448
;  PRIOR FILING DATE: 2000-02-29
;  NUMBER OF SEQ ID NOS: 2050
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1806
;   LENGTH: 1282
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(1282)
;   OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1806

```
  Query Match             39.8%;  Score 1147.8;  DB 9;  Length 1282;
  Best Local Similarity   97.7%;  Pred. No. 6.7e-276;
  Matches 1259;  Conservative    0;  Mismatches   18;  Indels   11;  Gaps    9;

Qy       1510 CCGAGGAGTTCGTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCA 1569
Db          1 CCGAGGAGTTCGTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCA 60

Qy       1570 CCAGCAACATGACCTATGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATA 1629
Db         61 CCAGCAACATGACCTATGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATA 120

Qy       1630 CAGGTATCAGCCTCCTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCT 1689
Db        121 CAGGAATCAGCCTCCTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCT 180

Qy       1690 ACCTCACGCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGC 1749
Db        181 ACCTCACGCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGC 240

Qy       1750 TGAGTCCCATCGTTAGCTGTGGACCCCCT-GGCGTCCTGCTCACCCGGCCAGTCATCCT- 1807
Db        241 TGAGTCCCATCGTTAGCTGTGGACCCCCTGGGCGTCCTGCTCACCCGGCCAGTCATCCTG 300

Qy       1808 GGCTATGGACCACTGTGGGGAGCCCAGCCCTGACAGCT-GGAGCCTGCGCCTCAAAAAGC 1866
Db        301 GGGTATGGACCACTGTGGGGAGCCCAGCCCTGACAGCTGGGAGCCTGCGCCTCAAAAAGC 360

Qy       1867 AGTCGTGCGAGGGCAGCTGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCC 1926
Db        361 AGTCGTGCGAGGGCAGCTGGGAGGATGTGCT--TGCACCTGGGCGAGGAGGCGCCCTCCC 418

Qy       1927 ACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCC 1986
Db        419 ACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGAGCC 478

Qy       1987 GCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTC 2046
Db        479 GCTATGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTC 538

Qy       2047 TGTTTGCGCCGGTGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATG 2106
Db        539 TGTTTGCGCCGGTGGCCTGCACCTCCCTCGAGTACAACATACTGGTCTACTGCCTGCATG 598

Qy       2107 ACACCCACGATGCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGA 2166
Db        599 ACACTCACGATGCACTCAACGTAGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGA 658

Qy       2167 TCCAGGAGCCACGGGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCC 2226
Db        659 TCCAGGAGCCACGGGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCC 718

Qy       2227 ACGATGTGCCCAGCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCT 2286
Db        719 ACGATGTGCCCAGCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCT 778

Qy       2287 TTTATCACATCTGGAATGGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTG 2346
Db        779 TTTATCACATCTGGAATGGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTG 838

Qy       2347 TCAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACG 2406
Db        839 TCAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACG 898

Qy       2407 GGCAGAGCTTCAGCATCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGG 2466
Db        899 GGCAGAGCTTCAGCATCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGG 958

Qy       2467 CTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCT 2526
Db        959 CTCTGGAGAGTGAAGCGNGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCT 1018

Qy       2527 TCCTCATTCGGCAGAAGATAA-TTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGAC 2585
Db       1019 TCCTCATTCGGCAGAAGATAATTTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGAC 1078

Qy       2586 TGGCGG-ACTCTGGCCCAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAA 2644
Db       1079 TGGCGGAACTCTGGCCCAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAA 1138

Qy       2645 GCCCAGCCCCACAGCCATGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAA 2704
Db       1139 GCCCAGCCCCACAGCCATGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAA 1198

Qy       2705 CCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCAC 2764
Db       1199 CCTCAGCCAGCTGG-TGCAGCAGTGGCTGGACTGGGC--AGCAGACGCTGG-CTCTTCAC 1254
```

```
Qy       2765 AGTGTCGGAGGCTGAGTGCTGAGGCCGG 2792
Db       1255 AGTGTCGGAGGCTGAGTGCTGAGGCCGG 1282
```

RESULT 12
US-10-087-684-1
; Sequence 1, Application US/10087684
; Publication No. US20040029116A1
; GENERAL INFORMATION:
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: MacDougall, John R.
; APPLICANT: Millet, Isabelle
; APPLICANT: Ellerman, Karen
; APPLICANT: Stone, David J.
; APPLICANT: Grosse, William M.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Cathereine E.
; APPLICANT: Casman, Stacie, J.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Li, Li
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Mishra, Vishnu
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Rastelli, Luca
; APPLICANT: Tchernev, Velizar T.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Guo, Xiaojia
; APPLICANT: Miller, Charles E.
; APPLICANT: Gangolli, Esha A.
; TITLE OF INVENTION: PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-214 CIP
; CURRENT APPLICATION NUMBER: US/10/087,684
; CURRENT FILING DATE: 2003-03-10
; PRIOR APPLICATION NUMBER: 60/253,834
; PRIOR FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: 60/250,926
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 60/264,180
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/274,194
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/313,656
; PRIOR FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 60/327,456
; PRIOR FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 220
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 1
; LENGTH: 2860
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (59)..(2857)
US-10-087-684-1

Query Match 30.9%; Score 889; DB 13; Length 2860;
Best Local Similarity 60.7%; Pred. No. 2.1e-211;
Matches 1637; Conservative 0; Mismatches 960; Indels 102; Gaps 7;

```
Qy        184 ACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATG 243
Db        168 ACTCCTTCCCGTCAGCGCCAGCAGAGCCGCTGCCCTACTTCCTGCAGGAGCCACAGGACG 227

Qy        244 TGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGA 303
Db        228 CCTACATTGTGAAGAACAAGCCTGTGGAGCTCCGCTGCCGCGCCTTCCCCGCCACACAGA 287

Qy        304 TCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCA 363
Db        288 TCTACTTCAAGTGCAACGGCGAGTGGGTCAGCCAGAACGACCACGTCACACAGGAAGGCC 347

Qy        364 CAGACGGGAGCAGTGGTGAGCCGACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGG 423
Db        348 TGGATGAGGCCACCGGTCTGCGGGTGCGCGAGGTGCAGATCGAGGTGTCGCGGCAGCAGG 407

Qy        424 TCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGG 483
Db        408 TGGAGGAGCTCTTTGGGCTGGAGGATTACTGGTGCCAGTGCGTGGCCTGGAGCTCCGCGG 467

Qy        484 GCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGC 543
Db        468 GCACCACCAAGAGTCGCCGAGCCTACGTCCGCATCGCCTACCTGCGCAAGAACTTCGATC 527

Qy        544 AGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCAC 603
Db        528 AGGAGCCTCTGGGCAAGGAGGTGCCCCTGGACCATGAGGTTCTCCTGCAGTGCCGCCCGC 587

Qy        604 CGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGT 663
Db        588 CGGAGGGGGTGCCTGTGGCCGAGGTGGAATGGCTCAAGAATGAGGATGTCATCGACCCCA 647

Qy        664 CCCTGGACCCCAATGTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCC 723
Db        648 CCCAGGACACCAACTTCCTGCTCACCATCGACCACAACCTCATCATCCGCCAGGCCCGCC 707

Qy        724 TTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCG 783
Db        708 TGTCGGACACTGCCAACTATACCTGCGTGGCCAAGAACATCGTGGCCAAACGCCGGAGCA 767

Qy        784 CCTCCGCTGCTGTCATCGTCTACGTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCG 843
Db        768 CCACTGCCACCGTCATCGTCTACGTGAATGGCGGCTGGTCCAGCTGGGCAGAGTGGTCAC 827

Qy        844 TCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGG 903
Db        828 CCTGCTCCAACCGCTGTGGCCGAGGCTGGCAGAAGCGCACCCGGACCTGCACCAACCCCG 887

Qy        904 CGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCCT 963
Db        888 CTCCACTCAACGGAGGGGCCTTCTGCGAGGGCCAGGCATTCCAGAA---GACCGCCTGCA 944

Qy        964 CTCTGCTTGTCTCTGTGGACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGC 1023
Db        945 CCACCATCTGCCCAGTCGATGGGGCGTGGACGGAGTGGAGCAAGTGGTCAGCCTGCAGCA 1004

Qy       1024 TGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGG 1083
Db       1005 CTGAGTGTGCCCACTGGCGTAGCCGCGAGTGCATGGCGCCCCCACCCCAGAACGGAGGCC 1064

Qy       1084 AGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGTACCAGTGACCTCTGTGTACACA 1143
Db       1065 GTGACTGCAGCGGGACGCTGCTCGACTCTAAGAACTGCACAGATGGGCTGTGCATGCAAC 1124

Qy       1144 GTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCC 1203
Db       1125 TGGAGGCCTCAGGGGATGCGGCGCTGTATGCGGGGCTCGTGGTGGCCATCTTCGTGGTCG 1184

Qy       1204 TGGTCCTGCTGCTGCTTGTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACT 1263
Db       1185 TGGCAATCCTCATGGCGGTGGGGGTGGTGGTGTACCGCCGCAACTGCCGTGACTTCGACA 1244

Qy       1264 CAGATGTGGCTGACTCGTCCATT---CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGC 1320
Db       1245 CAGACATCACTGACTCATCTGCTGCCCTGACTGGTGGTTTCCACCCCGTCAACTTTAAGA 1304

Qy       1321 CCAGCAAAGCAGACAACCCCCATCTGCT------CACCATCCAGCCGGACCTCAGCACCA 1374
Db       1305 CGGCAAGGCCCAGTAACCCGCAGCTCCTACACCCCTCTGTGCCTCCTGACCTGACAGCCA 1364

Qy       1375 CCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGC 1434
```

```
Db       1365 GCGCCGGCATCTACCGCGGACCCGTGTATGCCCTGCAGGACTCCACCGACAAAATCCCCA 1424
Qy       1435 TCACCAATGGGCACCTGCTCAGCCCC---------------------------------- 1460
Db       1425 TGACCAACTCTCCTCTGCTGGACCCCTTACCCAGCCTTAAGGTCAAGGTCTACAGCTCCA 1484
Qy       1461 --------------------------------------------CTGGGTGGCGGCCGCC 1476
Db       1485 GCACCACGGGCTCTGGGCCAGGCCTGGCAGATGGGGCTGACCTGCTGGGGGTCTTGCCGC 1544
Qy       1477 ACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCA 1536
Db       1545 CTGGCACATACCCTAGCGATTTCGCCCGGGACACCCACTTCCTGCACCTGCGCAGCGCCA 1604
Qy       1537 CCCAGAACTAC------TTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGA 1590
Db       1605 GCCTCGGTTCCCAGCAGCTCTTGGGCCTGCCCCGAGACCCAGGGAGCAGCGTCAGCGGCA 1664
Qy       1591 CCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCC 1650
Db       1665 CCTTTGGCTGCCTGGGTGGGAGGCTCAGCATCCCCGGCACAGGGGTCAGCTTGCTGGTGC 1724
Qy       1651 CCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGG 1710
Db       1725 CCAATGGAGCCATTCCCCAGGGCAAGTTCTACGAGATGTATCTACTCATCAACAAGGCAG 1784
Qy       1711 AAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTG 1770
Db       1785 AAAGTACCCTGCCGCTTTCAGAAGGGACCCAGACAGTATTGAGCCCCTCGGTGACCTGTG 1844
Qy       1771 GACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGC 1830
Db       1845 GACCCACAGGCCTCCTGCTGTGCCGCCCCGTCATCCTCACCATGCCCCACTGTGCCGAAG 1904
Qy       1831 CCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGC 1890
Db       1905 TCAGTGCCCGTGACTGGATCTTTCAGCTCAAGACCCAGGCCCACCAGGGCCACTGGGAG- 1963
Qy       1891 AGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGG 1950
Db       1964 --GAGGTGGTGACCCTGGATGAGGAGACCCTGAACACACCCTGCTACTGCCAGCTGGAGC 2021
Qy       1951 CCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCC 2010
Db       2022 CCAGGGCCTGTCACATCCTGCTGGACCAGCTGGGCACCTACGTGTTCACGGGCGAGTCCT 2081
Qy       2011 TCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCT 2070
Db       2082 ATTCCCGCTCAGCAGTCAAGCGGCTCCAGCTGGCCGTCTTCGCCCCCGCCCTCTGCACCT 2141
Qy       2071 CCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGG 2130
Db       2142 CCCTGGAGTACAGCCTCCGGGTCTACTGCCTGGAGGACACGCCTGTAGCACTGAAGGAGG 2201
Qy       2131 TGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACT 2190
Db       2202 TGCTGGAGCTGGAGCGGACTCTGGGCGGATACTTGGTGGAGGAGCCGAAACCGCTAATGT 2261
Qy       2191 TCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGA 2250
Db       2262 TCAAGGACAGTTACCACAACCTGCGCCTCTCCCTCCATGACCTCCCCCATGCCCATTGGA 2321
Qy       2251 AGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGC 2310
Db       2322 GGAGCAAGCTGCTGGCCAAATACCAGGAGATCCCCTTCTATCACATTTGGAGTGGCAGCC 2381
Qy       2311 AGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGG 2370
Db       2382 AGAAGGCCCTCCACTGCACTTTCACCCTGGAGAGGCACAGCTTGGCCTCCACAGAGCTCA 2441
Qy       2371 CCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCA 2430
Db       2442 CCTGCAAGATCTGCGTGCGGCAAGTGGAAGGGGAGGGCCAGATATTCCAGCTGCATACCA 2501
Qy       2431 ACATCACCAAG---GACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGG 2487
Db       2502 CTCTGGCAGAGACACCTGCTGGCTCCCTGGACACTCTCTGCTCTGCCCCTGGCAGCACTG 2561
Qy       2488 TCCCAGCCCTGGTGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAA 2547
Db       2562 TCACCACCCAGCTGGACCTTATGCCTTCAAGATCCCACTGTCCATCCGCCAGAAGATAT 2621
Qy       2548 TTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAAC 2607
Db       2622 GCAACAGCCTAGATGCCCCCAACTCACGGGGCAATGACTGGCGGATGTTAGCACAGAAGC 2681
Qy       2608 TCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCC 2667
Db       2682 TCTCTATGGACCGGTACCTGAATTACTTTGCCACCAAAGCGAGCCCCACGGGTGTGATCC 2741
Qy       2668 TCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAG 2727
Db       2742 TGGACCTCTGGGAAGCTCTGCAGCAGGACGATGGGGACCTCAACAGCCTGGCGAGTGCCT 2801
Qy       2728 TGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2786
Db       2802 TGGAGGAGATGGCAAGAGTGAGATGCTGGTGGCTGTGGCCACCGACGGGGACTGCTGA 2860
```

RESULT 13
US-10-218-779-1
; Sequence 1, Application US/10218779
; Publication No. US20040029222A1
; GENERAL INFORMATION:
;  APPLICANT: Edinger, Shlomit
;  APPLICANT:  MacDougall, John
;  APPLICANT:  Millet, Isabelle
;  APPLICANT:  Ellerman, Karen
;  APPLICANT:  Stone, David
;  APPLICANT:  Gerlach, Valerie
;  APPLICANT:  Grosse, William
;  APPLICANT:  Alsobrook II, John
;  APPLICANT:  Lepley, Denise
;  APPLICANT:  Rieger, Daniel
;  APPLICANT:  Burgess, Catherine
;  APPLICANT:  Casman, Stacie
;  APPLICANT:  Spytek, Kimberly
;  APPLICANT:  Boldog, Ferenc
;  APPLICANT:  Li, Li
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Mishra, Vishnu
;  APPLICANT:  Patturajan, Meera
;  APPLICANT:  Shenoy, Suresh
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Tchernev, Velizar
;  APPLICANT:  Vernet, Corine
;  APPLICANT:  Zerhusen, Bryan
;  APPLICANT:  Malyankar, Uriel
;  APPLICANT:  Guo, Xiaojia
;  APPLICANT:  Miller, Charles
;  APPLICANT:  Gangolli, Esha
;  TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
;  FILE REFERENCE: 21402-214
;  CURRENT APPLICATION NUMBER: US/10/218,779
;  CURRENT FILING DATE:  2002-08-14
;  PRIOR APPLICATION NUMBER: 60/253,834
;  PRIOR FILING DATE: 2000-11-29
;  PRIOR APPLICATION NUMBER: 60/250,-926
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: 60/264,180
;  PRIOR FILING DATE: 2001-01-25
;  PRIOR APPLICATION NUMBER: 60/313,656
;  PRIOR FILING DATE: 2001-08-20
;  PRIOR APPLICATION NUMBER: 60/327,456
;  PRIOR FILING DATE: 2001-10-05

```
;  NUMBER OF SEQ ID NOS: 216
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 2860
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-218-779-1

  Query Match             30.9%;  Score 889;  DB 13;  Length 2860;
  Best Local Similarity   60.7%;  Pred. No. 2.1e-211;
  Matches 1637;  Conservative    0;  Mismatches  960;  Indels  102;  Gaps    7;

Qy        184 ACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATG 243
Db        168 ACTCCTTCCCGTCAGCGCCAGCAGAGCCGCTGCCCTACTTCCTGCAGGAGCCACAGGACG 227

Qy        244 TGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGA 303
Db        228 CCTACATTGTGAAGAACAAGCCTGTGGAGCTCCGCTGCCGCGCCTTCCCCGCCACACAGA 287

Qy        304 TCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCA 363
Db        288 TCTACTTCAAGTGCAACGGCGAGTGGGTCAGCCAGAACGACCACGTCACACAGGAAGGCC 347

Qy        364 CAGACGGGAGCAGTGGTGAGCCGACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGG 423
Db        348 TGGATGAGGCCACCGGTCTGCGGGTGCGCGAGGTGCAGATCGAGGTGTCGCGGCAGCAGG 407

Qy        424 TCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGG 483
Db        408 TGGAGGAGCTCTTTGGGCTGGAGGATTACTGGTGCCAGTGCGTGGCCTGGAGCTCCGCGG 467

Qy        484 GCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGC 543
Db        468 GCACCACCAAGAGTCGCCGAGCCTACGTCCGCATCGCCTACCTGCGCAAGAACTTCGATC 527

Qy        544 AGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCAC 603
Db        528 AGGAGCCTCTGGGCAAGGAGGTGCCCCTGGACCATGAGGTTCTCCTGCAGTGCCGCCCGC 587

Qy        604 CGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGT 663
Db        588 CGGAGGGGGTGCCTGTGGCCGAGGTGGAATGGCTCAAGAATGAGGATGTCATCGACCCCA 647

Qy        664 CCCTGGACCCCAATGTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCC 723
Db        648 CCCAGGACACCAACTTCCTGCTCACCATCGACCACAACCTCATCATCCGCCAGGCCCGCC 707

Qy        724 TTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCG 783
Db        708 TGTCGGACACTGCCAACTATACCTGCGTGGCCAAGAACATCGTGGCCAAACGCCGGAGCA 767

Qy        784 CCTCCGCTGCTGTCATCGTCTACGTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCG 843
Db        768 CCACTGCCACCGTCATCGTCTACGTGAATGGCGGCTGGTCCAGCTGGGCAGAGTGGTCAC 827

Qy        844 TCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGG 903
Db        828 CCTGCTCCAACCGCTGTGGCCGAGGCTGGCAGAAGCGCACCCGGACCTGCACCAACCCCG 887

Qy        904 CGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCCT 963
Db        888 CTCCACTCAACGGAGGGGCCTTCTGCGAGGGCCAGGCATTCCAGAA---GACCGCCTGCA 944

Qy        964 CTCTGCTTGTCTCTGTGGACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGC 1023
Db        945 CCACCATCTGCCCAGTCGATGGGGCGTGGACGGAGTGGAGCAAGTGGTCAGCCTGCAGCA 1004

Qy       1024 TGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGG 1083
Db       1005 CTGAGTGTGCCCACTGGCGTAGCCGCGAGTGCATGGCGCCCCCACCCCAGAACGGAGGCC 1064

Qy       1084 AGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGTACCAGTGACCTCTGTGTACACA 1143
Db       1065 GTGACTGCAGCGGGACGCTGCTCGACTCTAAGAACTGCACAGATGGGCTGTGCATGCAAC 1124

Qy       1144 GTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCC 1203
Db       1125 TGGAGGCCTCAGGGGATGCGGCGCTGTATGCGGGGCTCGTGGTGGCCATCTTCGTGGTCG 1184

Qy       1204 TGGTCCTGCTGCTGCTTGTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACT 1263
Db       1185 TGGCAATCCTCATGGCGGTGGGGGTGGTGGTGTACCGCCGCAACTGCCGTGACTTCGACA 1244

Qy       1264 CAGATGTGGCTGACTCGTCCATT---CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGC 1320
Db       1245 CAGACATCACTGACTCATCTGCTGCCCTGACTGGTGGTTTCCACCCCGTCAACTTTAAGA 1304

Qy       1321 CCAGCAAAGCAGACAACCCCCATCTGCT------CACCATCCAGCCGGACCTCAGCACCA 1374
Db       1305 CGGCAAGGCCCAGTAACCCGCAGCTCCTACACCCCTCTGTGCCTCCTGACCTGACAGCCA 1364

Qy       1375 CCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGC 1434
Db       1365 GCGCCGGCATCTACCGCGGACCCGTGTATGCCCTGCAGGACTCCACCGACAAAATCCCCA 1424

Qy       1435 TCACCAATGGGCACCTGCTCAGCCCC---------------------------------- 1460
Db       1425 TGACCAACTCTCCTCTGCTGGACCCCTTACCCAGCCTTAAGGTCAAGGTCTACAGCTCCA 1484

Qy       1461 ------------------------------------------CTGGGTGGCGGCCGCC 1476
Db       1485 GCACCACGGGCTCTGGGCCAGGCCTGGCAGATGGGGCTGACCTGCTGGGGGTCTTGCCGC 1544

Qy       1477 ACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCA 1536
Db       1545 CTGGCACATACCCTAGCGATTTCGCCCGGGACACCCACTTCCTGCACCTGCGCAGCGCCA 1604

Qy       1537 CCCAGAACTAC------TTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGA 1590
Db       1605 GCCTCGGTTCCCAGCAGCTCTTGGGCCTGCCCCGAGACCCAGGGAGCAGCGTCAGCGGCA 1664

Qy       1591 CCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCC 1650
Db       1665 CCTTTGGCTGCCTGGGTGGGAGGCTCAGCATCCCCGGCACAGGGGTCAGCTTGCTGGTGC 1724

Qy       1651 CCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGG 1710
Db       1725 CCAATGGAGCCATTCCCCAGGGCAAGTTCTACGAGATGTATCTACTCATCAACAAGGCAG 1784

Qy       1711 AAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTG 1770
Db       1785 AAAGTACCCTGCCGCTTTCAGAAGGGACCCAGACAGTATTGAGCCCCTCGGTGACCTGTG 1844

Qy       1771 GACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGC 1830
Db       1845 GACCCACAGGCCTCCTGCTGTGCCGCCCCGTCATCCTCACCATGCCCCACTGTGCCGAAG 1904

Qy       1831 CCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGC 1890
Db       1905 TCAGTGCCCGTGACTGGATCTTTCAGCTCAAGACCCAGGCCCACCAGGGCCACTGGGAG- 1963

Qy       1891 AGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGG 1950
Db       1964 --GAGGTGGTGACCCTGGATGAGGAGACCCTGAACACACCCTGCTACTGCCAGCTGGAGC 2021

Qy       1951 CCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCC 2010
Db       2022 CCAGGGCCTGTCACATCCTGCTGGACCAGCTGGGCACCTACGTGTTCACGGGCGAGTCCT 2081

Qy       2011 TCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCT 2070
Db       2082 ATTCCCGCTCAGCAGTCAAGCGGCTCCAGCTGGCCGTCTTCGCCCCCGCCCTCTGCACCT 2141

Qy       2071 CCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGG 2130
```

```
Db    2142 CCCTGGAGTACAGCCTCCGGGTCTACTGCCTGGAGGACACGCCTGTAGCACTGAAGGAGG 2201

Qy    2131 TGGTGCAGCTGGAGAAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACT 2190
Db    2202 TGCTGGAGCTGGAGCGGACTCTGGGCGGATACTTGGTGGAGGAGCCGAAACCGCTAATGT 2261

Qy    2191 TCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGA 2250
Db    2262 TCAAGGACAGTTACCACAACCTGCGCCTCTCCCTCCATGACCTCCCCCATGCCCATTGGA 2321

Qy    2251 AGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGC 2310
Db    2322 GGAGCAAGCTGCTGGCCAAATACCAGGAGATCCCCTTCTATCACATTTGGAGTGGCAGCC 2381

Qy    2311 AGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGG 2370
Db    2382 AGAAGGCCCTCCACTGCACTTTCACCCTGGAGAGGCACAGCTTGGCCTCCACAGAGCTCA 2441

Qy    2371 CCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCA 2430
Db    2442 CCTGCAAGATCTGCGTGCGGCAAGTGGAAGGGGAGGGCCAGATATTCCAGCTGCATACCA 2501

Qy    2431 ACATCACCAAG---GACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGG 2487
Db    2502 CTCTGGCAGAGACACCTGCTGGCTCCCTGGACACTCTCTGCTCTGCCCCTGGCAGCACTG 2561

Qy    2488 TCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAA 2547
Db    2562 TCACCACCCAGCTGGGACCTTATGCCTTCAAGATCCCACTGTCCATCCGCCAGAAGATAT 2621

Qy    2548 TTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAAC 2607
Db    2622 GCAACAGCCTAGATGCCCCCAACTCACGGGGCAATGACTGGCGGATGTTAGCACAGAAGC 2681

Qy    2608 TCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCC 2667
Db    2682 TCTCTATGGACCGGTACCTGAATTACTTTGCCACCAAAGCGAGCCCCACGGGTGTGATCC 2741

Qy    2668 TCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAG 2727
Db    2742 TGGACCTCTGGGAAGCTCTGCAGCAGGACGATGGGGACCTCAACAGCCTGGCGAGTGCCT 2801

Qy    2728 TGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2786
Db    2802 TGGAGGAGATGGGCAAGAGTGAGATGCTGGTGGCTGTGGCCACCGACGGGGACTGCTGA 2860
```

RESULT 14
US-10-087-684-3
; Sequence 3, Application US/10087684
; Publication No. US20040029116A1
; GENERAL INFORMATION:
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: MacDougall, John R.
; APPLICANT: Millet, Isabelle
; APPLICANT: Ellerman, Karen
; APPLICANT: Stone, David J.
; APPLICANT: Grosse, William M.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Cathereine E.
; APPLICANT: Casman, Stacie, J.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Li, Li
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Mishra, Vishnu
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Rastelli, Luca
; APPLICANT: Tchernev, Velizar T.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Guo, Xiaojia
; APPLICANT: Miller, Charles E.
; APPLICANT: Gangolli, Esha A.
; TITLE OF INVENTION: PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-214 CIP
; CURRENT APPLICATION NUMBER: US/10/087,684
; CURRENT FILING DATE: 2003-03-10
; PRIOR APPLICATION NUMBER: 60/253,834
; PRIOR FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: 60/250,926
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 60/264,180
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/274,194
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/313,656
; PRIOR FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 60/327,456
; PRIOR FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 220
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 3
;  LENGTH: 2860
;  TYPE: DNA
;  ORGANISM: Homo sapiens
;  FEATURE:
;  NAME/KEY: CDS
;  LOCATION: (59)..(2857)
US-10-087-684-3

```
  Query Match             30.8%;  Score 887.4;  DB 13;  Length 2860;
  Best Local Similarity   60.6%;  Pred. No. 5.4e-211;
  Matches 1636;  Conservative    0;  Mismatches  961;  Indels  102;  Gaps    7;

Qy     184 ACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATG 243
Db     168 ACTCCTTCCCGTCAGCGCCAGCAGAGCCGCTGCCCTACTTCCTGCAGGAGCCACAGGACG 227

Qy     244 TGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGA 303
Db     228 CCTACATTGTGAAGAACAAGCCTGTGGAGCTTCGCTGCCGCGCCTTCCCCGCCACACAGA 287

Qy     304 TCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCA 363
Db     288 TCTACTTCAAGTGCAACGGCGAGTGGGTCAGCCAGAACGACCACGTCACACAGGAAGGCC 347

Qy     364 CAGACGGGAGCAGTGGTGAGCCGACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGG 423
Db     348 TGGATGAGGCCACCGGCCTGCGGGTGCGCGAGGTGCAGATCGAGGTGTCGCGGCAGCAGG 407

Qy     424 TCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGG 483
Db     408 TGGAGGAGCTCTTTGGGCTGGAGGATTACTGGTGCCAGTGCGTGGCCTGGAGCTCCGCAG 467

Qy     484 GCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGC 543
Db     468 GCACCACCAAGAGTCGCCGAGCCTACGTCCGCATCGCCTACCTGCGCAAGAACTTCGATC 527

Qy     544 AGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCAC 603
Db     528 AGGAGCCTCTGGGCAAGGAGGTGCCCCTGGACCATGAGGTTCTCCTGCAGTGCCGCCCGC 587

Qy     604 CGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGT 663
Db     588 CGGAGGGGGTGCCTGTGGCCGAGGTGGAATGGCTCAAGAATGAGGATGTCATCGACCCCA 647

Qy     664 CCCTGGACCCCAATGTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCC 723
Db     648 CCCAGGACACCAACTTCCTGCTCACCATCGACCACAACCTCATCATCCGCCAGGCCCGCC 707

Qy     724 TTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCG 783
Db     708 TGTCGGACACTGCCAACTATACCTGCGTGGCCAAGAACATCGTGGCCAAACGCCGGAGCA 767
```

```
Qy    784 CCTCCGCTGCTGTCATCGTCTACGTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCG 843
Db    768 CCACTGCCACCGTCATCGTCTACGTGAATGGCGGCTGGTCCAGCTGGGCAGAGTGGTCAC 827

Qy    844 TCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGG 903
Db    828 CCTGCTCCAACCGCTGTGGCCGAGGCTGGCAGAAGCGCACCCGGACCTGCACCAACCCCG 887

Qy    904 CGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCCT 963
Db    888 CTCCACTCAACGGAGGGGCCTTCTGCGAGGGCCAGGCATTCCAGAA---GACCGCCTGCA 944

Qy    964 CTCTGCTTGTCTCTGTGGACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGC 1023
Db    945 CCACCATCTGCCCAGTCGATGGGGCGTGGACGGAGTGGAGCAAGTGGTCAGCCTGCAGCA 1004

Qy   1024 TGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGG 1083
Db   1005 CTGAGTGTGCCCACTGGCGTAGCCGCGAGTGCATGGCGCCCCCACCCCAGAACGGAGGCC 1064

Qy   1084 AGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGTACCAGTGACCTCTGTGTACACA 1143
Db   1065 GTGACTGCAGCGGGACGCTGCTCGACTCTAAGAACTGCACAGATGGGCTGTGCATGCAAC 1124

Qy   1144 GTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCC 1203
Db   1125 TGGAGGCCTCAGGGGATGCGGCGCTGTATGCGGGGCTCGTGGTGGCCATCTTCGTGGTCG 1184

Qy   1204 TGGTCCTGCTGCTGCTTGTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACT 1263
Db   1185 TGGCAATCCTCATGGCGGTGGGGGTGGTGGTGTACCGCCGCAACTGCCGTGACTTCGACA 1244

Qy   1264 CAGATGTGGCTGACTCGTCCATT---CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGC 1320
Db   1245 CAGACATCACTGACTCATCTGCTGCCCTGACTGGTGGTTTCCACCCCGTCAACTTTAAGA 1304

Qy   1321 CCAGCAAAGCAGACAACCCCCATCTGCT------CACCATCCAGCCGGACCTCAGCACCA 1374
Db   1305 CGGCAAGGCCCAGTAACCCGCAGCTCCTACACCCCTCTGTGCCTCCTGACCTGACAGCCA 1364

Qy   1375 CCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGC 1434
Db   1365 GCGCCGGCATCTACCGCGGACCCGTGTATGCCCTGCAGGACTCCACCGACAAAATCCCCA 1424

Qy   1435 TCACCAATGGGCACCTGCTCAGCCCC-------------------------------- 1460
Db   1425 TGACCAACTCTCCTCTGCTGGACCCCTTACCCAGCCTTAAGGTCAAGGTCTACAGCTCCA 1484

Qy   1461 ------------------------------------------CTGGGTGGCGGCCGCC 1476
Db   1485 GCACCACGGGCTCTGGGCCAGGCCTGGCAGATGGGGCTGACCTGCTGGGGGTCTTGCCGC 1544

Qy   1477 ACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCA 1536
Db   1545 CTGGCACATACCCTAGCGATTTCGCCCGGGACACCCACTTCCTGCACCTGCGCAGCGCCA 1604

Qy   1537 CCCAGAACTAC------TTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGA 1590
Db   1605 GCCTCGGTTCCCAGCAGCTCTTGGGCCTGCCCCGAGACCCAGGGAGCAGCGTCAGCGGCA 1664

Qy   1591 CCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCC 1650
Db   1665 CCTTTGGCTGCCTGGGTGGAGGCTCAGCATCCCCGGCACAGGGGTCAGCTTGCTGGTGC 1724

Qy   1651 CCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGG 1710
Db   1725 CCAATGGAGCCATTCCCCAGGGCAAGTTCTACGAGATGTATCTACTCATCAACAAGGCAG 1784

Qy   1711 AAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTG 1770
Db   1785 AAAGTACCCTGCCGCTTTCAGAAGGGACCCAGACAGTATTGAGCCCCTCGGTGACCTGTG 1844

Qy   1771 GACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGC 1830
Db   1845 GACCCACAGGCCTCCTGCTGTGCCGCCCCGTCATCCTCACCATGCCCCACTGTGCCGAAG 1904

Qy   1831 CCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGC 1890
Db   1905 TCAGTGCCCGTGACTGGATCTTTCAGCTCAAGACCCAGGCCCACCAGGGCCACTGGGAG- 1963

Qy   1891 AGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGG 1950
Db   1964 --GAGGTGGTGACCCTGGATGAGGAGACCCTGAACACACCCTGCTACTGCCAGCTGGAGC 2021

Qy   1951 CCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCC 2010
Db   2022 CCAGGGCCTGTCACATCCTGCTGGACCAGCTGGGCACCTACGTGTTCACGGGCGAGTCCT 2081

Qy   2011 TCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCT 2070
Db   2082 ATTCCCGCTCAGCAGTCAAGCGGCTCCAGCTGGCCGTCTTCGCCCCCGCCCTCTGCACCT 2141

Qy   2071 CCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGG 2130
Db   2142 CCCTGGAGTACAGCCTCCGGGTCTACTGCCTGGAGGACACGCCTGTAGCACTGAAGGAGG 2201

Qy   2131 TGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACT 2190
Db   2202 TGCTGGAGCTGGAGCGGACTCTGGGCGGATACTTGGTGGAGGAGCCGAAACCGCTAATGT 2261

Qy   2191 TCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGA 2250
Db   2262 TCAAGGACAGTTACCACAACCTGCGCCTCTCCCTCCATGACCTCCCCCATGCCCATTGGA 2321

Qy   2251 AGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGC 2310
Db   2322 GGAGCAAGCTGCTGGCCAAATACCAGGAGATCCCCTTCTATCACATTTGGAGTGGCAGCC 2381

Qy   2311 AGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGG 2370
Db   2382 AGAAGGCCCTCCACTGCACTTTCACCCTGGAGAGGCACAGCTTGGCCTCCACAGAGCTCA 2441

Qy   2371 CCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCA 2430
Db   2442 CCTGCAAGATCTGCGTGCGGCAAGTGGAAGGGGAGGGCCAGATATTCCAGCTGCATACCA 2501

Qy   2431 ACATCACCAAG---GACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGG 2487
Db   2502 CTCTGGCAGAGACACCTGCTGGCTCCCTGGACACTCTCTGCTCTGCCCCTGGCAGCACTG 2561

Qy   2488 TCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAA 2547
Db   2562 TCACCACCCAGCTGGGACCTTATGCCTTCAAGATCCCACTGTCCATCCGCCAGAAGATAT 2621

Qy   2548 TTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAAC 2607
Db   2622 GCAACAGCCTAGATGCCCCCAACTCACGGGGCAATGACTGGCGGATGTTAGCACAGAAGC 2681

Qy   2608 TCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCC 2667
Db   2682 TCTCTATGGACCGGTACCTGAATTACTTTGCCACCAAAGCGAGCCCCACGGGTGTGATCC 2741

Qy   2668 TCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAG 2727
Db   2742 TGGACCTCTGGGAAGCTCTGCAGCAGGACGATGGGGACCTCAACAGCCTGGCGAGTGCCT 2801

Qy   2728 TGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2786
Db   2802 TGGAGGAGATGGGCAAGAGTGAGATGCTGGTGGCTGTGGCCACCGACGGGGACTGCTGA 2860
```

RESULT 15
US-10-218-779-3
; Sequence 3, Application US/10218779
; Publication No. US20040029222A1

```
; GENERAL INFORMATION:
;  APPLICANT: Edinger, Shlomit
;  APPLICANT:  MacDougall, John
;  APPLICANT:  Millet, Isabelle
;  APPLICANT:  Ellerman, Karen
;  APPLICANT:  Stone, David
;  APPLICANT:  Gerlach, Valerie
;  APPLICANT:  Grosse, William
;  APPLICANT:  Alsobrook II, John
;  APPLICANT:  Lepley, Denise
;  APPLICANT:  Rieger, Daniel
;  APPLICANT:  Burgess, Catherine
;  APPLICANT:  Casman, Stacie
;  APPLICANT:  Spytek, Kimberly
;  APPLICANT:  Boldog, Ferenc
;  APPLICANT:  Li, Li
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Mishra, Vishnu
;  APPLICANT:  Patturajan, Meera
;  APPLICANT:  Shenoy, Suresh
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Tchernev, Velizar
;  APPLICANT:  Vernet, Corine
;  APPLICANT:  Zerhusen, Bryan
;  APPLICANT:  Malyankar, Uriel
;  APPLICANT:  Guo, Xiaojia
;  APPLICANT:  Miller, Charles
;  APPLICANT:  Gangolli, Esha
;  TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
;  FILE REFERENCE: 21402-214
;  CURRENT APPLICATION NUMBER: US/10/218,779
;  CURRENT FILING DATE:  2002-08-14
;  PRIOR APPLICATION NUMBER: 60/253,834
;  PRIOR FILING DATE: 2000-11-29
;  PRIOR APPLICATION NUMBER: 60/250,-926
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: 60/264,180
;  PRIOR FILING DATE: 2001-01-25
;  PRIOR APPLICATION NUMBER: 60/313,656
;  PRIOR FILING DATE: 2001-08-20
;  PRIOR APPLICATION NUMBER: 60/327,456
;  PRIOR FILING DATE: 2001-10-05
;  NUMBER OF SEQ ID NOS: 216
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 2860
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-218-779-3

  Query Match             30.8%;  Score 887.4;  DB 13;  Length 2860;
  Best Local Similarity   60.6%;  Pred. No. 5.4e-211;
  Matches 1636;  Conservative    0;  Mismatches  961;  Indels  102;  Gaps    7;

Qy        184 ACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATG 243
Db        168 ACTCCTTCCCGTCAGCGCCAGCAGAGCCGCTGCCCTACTTCCTGCAGGAGCCACAGGACG 227

Qy        244 TGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGA 303
Db        228 CCTACATTGTGAAGAACAAGCCTGTGGAGCTTCGCTGCCGCGCCTTCCCCGCCACACAGA 287

Qy        304 TCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCA 363
Db        288 TCTACTTCAAGTGCAACGGCGAGTGGGTCAGCCAGAACGACCACGTCACACAGGAAGGCC 347

Qy        364 CAGACGGGAGCAGTGGTGAGCCGACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGG 423
Db        348 TGGATGAGGCCACCGGCCTGCGGGTGCGCGAGGTGCAGATCGAGGTGTCGCGGCAGCAGG 407

Qy        424 TCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGG 483
Db        408 TGGAGGAGCTCTTTGGGCTGGAGGATTACTGGTGCCAGTGCGTGGCCTGGAGCTCCGCAG 467

Qy        484 GCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGC 543
Db        468 GCACCACCAAGAGTCGCCGAGCCTACGTCCGCATCGCCTACCTGCGCAAGAACTTCGATC 527

Qy        544 AGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCAC 603
Db        528 AGGAGCCTCTGGGCAAGGAGGTGCCCCTGGACCATGAGGTTCTCCTGCAGTGCCGCCCGC 587

Qy        604 CGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGT 663
Db        588 CGGAGGGGGTGCCTGTGGCCGAGGTGGAATGGCTCAAGAATGAGGATGTCATCGACCCCA 647

Qy        664 CCCTGGACCCCAATGTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCC 723
Db        648 CCCAGGACACCAACTTCCTGCTCACCATCGACCACAACCTCATCATCCGCCAGGCCCGCC 707

Qy        724 TTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCG 783
Db        708 TGTCGGACACTGCCAACTATACCTGCGTGGCCAAGAACATCGTGGCCAAACGCCGGAGCA 767

Qy        784 CCTCCGCTGCTGTCATCGTCTACGTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCG 843
Db        768 CCACTGCCACCGTCATCGTCTACGTGAATGGCGGCTGGTCCAGCTGGGCAGAGTGGTCAC 827

Qy        844 TCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGG 903
Db        828 CCTGCTCCAACCGCTGTGGCCGAGGCTGGCAGAAGCGCACCCGGACCTGCACCAACCCCG 887

Qy        904 CGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCCT 963
Db        888 CTCCACTCAACGGAGGGGCCTTCTGCGAGGGCCAGGCATTCCAGAA---GACCGCCTGCA 944

Qy        964 CTCTGCTTGTCTCTGTGGACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGC 1023
Db        945 CCACCATCTGCCCAGTCGATGGGGCGTGGACGGAGTGGAGCAAGTGGTCAGCCTGCAGCA 1004

Qy       1024 TGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGG 1083
Db       1005 CTGAGTGTGCCCACTGGCGTAGCCGCGAGTGCATGGCGCCCCCACCCCAGAACGGAGGCC 1064

Qy       1084 AGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGTACCAGTGACCTCTGTGTACACA 1143
Db       1065 GTGACTGCAGCGGGACGCTGCTCGACTCTAAGAACTGCACAGATGGGCTGTGCATGCAAC 1124

Qy       1144 GTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCC 1203
Db       1125 TGGAGGCCTCAGGGGATGCGGCGCTGTATGCGGGGCTCGTGGTGGCCATCTTCGTGGTCG 1184

Qy       1204 TGGTCCTGCTGCTGCTTGTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACT 1263
Db       1185 TGGCAATCCTCATGGCGGTGGGGGTGGTGGTGTACCGCCGCAACTGCCGTGACTTCGACA 1244

Qy       1264 CAGATGTGGCTGACTCGTCCATT---CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGC 1320
Db       1245 CAGACATCACTGACTCATCTGCTGCCCTGACTGGTGGTTTCCACCCCGTCAACTTTAAGA 1304

Qy       1321 CCAGCAAAGCAGACAACCCCCATCTGCT------CACCATCCAGCCGGACCTCAGCACCA 1374
Db       1305 CGGCAAGGCCCAGTAACCCGCAGCTCCTACACCCCTCTGTGCCTCCTGACCTGACAGCCA 1364

Qy       1375 CCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGC 1434
Db       1365 GCGCCGGCATCTACCGCGGACCCGTGTATGCCCTGCAGGACTCCACCGACAAAATCCCCA 1424

Qy       1435 TCACCAATGGGCACCTGCTCAGCCCC---------------------------------- 1460
Db       1425 TGACCAACTCTCCTCTGCTGGACCCCTTACCCAGCCTTAAGGTCAAGGTCTACAGCTCCA 1484

Qy       1461 ------------------------------------------CTGGGTGGCGGCCGCC 1476
Db       1485 GCACCACGGGCTCTGGGCCAGGCCTGGCAGATGGGGCTGACCTGCTGGGGGTCTTGCCGC 1544
```

```
Qy   1477 ACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCA 1536
Db   1545 CTGGCACATACCCTAGCGATTTCGCCCGGGACACCCACTTCCTGCACCTGCGCAGCGCCA 1604

Qy   1537 CCCAGAACTAC------TTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGA 1590
Db   1605 GCCTCGGTTCCCAGCAGCTCTTGGGCCTGCCCCGAGACCCAGGGAGCAGCGTCAGCGGCA 1664

Qy   1591 CCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCC 1650
Db   1665 CCTTTGGCTGCCTGGGTGGAGGCTCAGCATCCCCGGCACAGGGGTCAGCTTGCTGGTGC 1724

Qy   1651 CCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGG 1710
Db   1725 CCAATGGAGCCATTCCCCAGGGCAAGTTCTACGAGATGTATCTACTCATCAACAAGGCAG 1784

Qy   1711 AAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTG 1770
Db   1785 AAAGTACCCTGCCGCTTTCAGAAGGGACCCAGACAGTATTGAGCCCCTCGGTGACCTGTG 1844

Qy   1771 GACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGC 1830
Db   1845 GACCCACAGGCCTCCTGCTGTGCCGCCCCGTCATCCTCACCATGCCCCACTGTGCCGAAG 1904

Qy   1831 CCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGC 1890
Db   1905 TCAGTGCCCGTGACTGGATCTTTCAGCTCAAGACCCAGGCCCACCAGGGCCACTGGGAG- 1963

Qy   1891 AGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGG 1950
Db   1964 --GAGGTGGTGACCCTGGATGAGGAGACCCTGAACACACCCTGCTACTGCCAGCTGGAGC 2021

Qy   1951 CCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCC 2010
Db   2022 CCAGGGCCTGTCACATCCTGCTGGACCAGCTGGGCACCTACGTGTTCACGGGCGAGTCCT 2081

Qy   2011 TCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCT 2070
Db   2082 ATTCCCGCTCAGCAGTCAAGCGGCTCCAGCTGGCCGTCTTCGCCCCCGCCCTCTGCACCT 2141

Qy   2071 CCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGG 2130
Db   2142 CCCTGGAGTACAGCCTCCGGGTCTACTGCCTGGAGGACACGCCTGTAGCACTGAAGGAGG 2201

Qy   2131 TGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACT 2190
Db   2202 TGCTGGAGCTGGAGCGGACTCTGGGCGGATACTTGGTGGAGGAGCCGAAACCGCTAATGT 2261

Qy   2191 TCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGA 2250
Db   2262 TCAAGGACAGTTACCACAACCTGCGCCTCTCCCTCCATGACCTCCCCCATGCCCATTGGA 2321

Qy   2251 AGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGC 2310
Db   2322 GGAGCAAGCTGCTGGCCAAATACCAGGAGATCCCCTTCTATCACATTTGGAGTGGCAGCC 2381

Qy   2311 AGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGG 2370
Db   2382 AGAAGGCCCTCCACTGCACTTTCACCCTGGAGAGGCACAGCTTGGCCTCCACAGAGCTCA 2441

Qy   2371 CCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCA 2430
Db   2442 CCTGCAAGATCTGCGTGCGGCAAGTGGAAGGGGAGGGCCAGATATTCCAGCTGCATACCA 2501

Qy   2431 ACATCACCAAG---GACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGG 2487
Db   2502 CTCTGGCAGAGACACCTGCTGGCTCCCTGGACACTCTCTGCTCTGCCCCTGGCAGCACTG 2561

Qy   2488 TCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAA 2547
Db   2562 TCACCACCCAGCTGGGACCTTATGCCTTCAAGATCCCACTGTCCATCCGCCAGAAGATAT 2621

Qy   2548 TTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAAC 2607
Db   2622 GCAACAGCCTAGATGCCCCCAACTCACGGGGCAATGACTGGCGGATGTTAGCACAGAAGC 2681

Qy   2608 TCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCC 2667
Db   2682 TCTCTATGGACCGGTACCTGAATTACTTTGCCACCAAAGCGAGCCCCACGGGTGTGATCC 2741

Qy   2668 TCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAG 2727
Db   2742 TGGACCTCTGGGAAGCTCTGCAGCAGGACGATGGGGACCTCAACAGCCTGGCGAGTGCCT 2801

Qy   2728 TGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2786
Db   2802 TGGAGGAGATGGGCAAGAGTGAGATGCTGGTGGCTGTGGCCACCGACGGGGACTGCTGA 2860
```

Search completed: October 6, 2004, 18:16:15
Job time : 1359 secs

```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2004  Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:        October  6, 2004, 11:46:57 ; Search time 7179 Seconds
                                            (without alignments)
                                            11983.973 Million cell updates/sec

Title:          US-09-970-944-1
Perfect score:  2881
Sequence:       1 agctggggctccgggctgag..........ccttccccacaccggggaga 2881

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters:      55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      EST:*
                1:  em_estba:*
                2:  em_esthum:*
                3:  em_estin:*
                4:  em_estmu:*
                5:  em_estov:*
                6:  em_estpl:*
                7:  em_estro:*
                8:  em_htc:*
                9:  gb_est1:*
                10:  gb_est2:*
                11:  gb_htc:*
                12:  gb_est3:*
                13:  gb_est4:*
                14:  gb_est5:*
                15:  em_estfun:*
                16:  em_estom:*
                17:  em_gss_hum:*
                18:  em_gss_inv:*
                19:  em_gss_pln:*
                20:  em_gss_vrt:*
                21:  em_gss_fun:*
                22:  em_gss_mam:*
                23:  em_gss_mus:*
                24:  em_gss_pro:*
                25:  em_gss_rod:*
                26:  em_gss_phg:*
                27:  em_gss_vrl:*
                28:  gb_gss1:*
                29:  gb_gss2:*

    Pred. No. is the number of results predicted by chance to have a
    score greater than or equal to the score of the result being printed,
    and is derived by analysis of the total score distribution.

                              SUMMARIES

                   %
Result           Query
  No.   Score    Match Length DB  ID                      Description
-----------------------------------------------------------------------
    1   907.8     31.5   2802  29  AY406491                AY406491 Homo sapi
    2   898.6     31.2   2791  29  AY406493                AY406493 Mus muscu
    3   871.6     30.3   3790  11  AK031655                AK031655 Mus muscu
    4   818.4     28.4   3866  11  AK018177                AK018177 Mus muscu
    5   763.2     26.5   2532  29  AY411747                AY411747 Homo sapi
    6   763.2     26.5   2802  29  AY406492                AY406492 Pan trogl
    7   756.2     26.2   1034  12  BI758231                BI758231 603029876
    8   726.4     25.2   2532  29  AY411749                AY411749 Mus muscu
    9   726.2     25.2    909  13  BX364574                BX364574 BX364574
   10   722.2     25.1    788   9  AI951556                AI951556 wv36f04.x
   11   711.2     24.7   1532  11  BC033727                BC033727 Homo sapi
   12     710     24.6    818  12  BI818609                BI818609 603033362
c  13   709.8     24.6    843  13  BX348193                BX348193 BX348193
   14   700.2     24.3    939  13  BX345406                BX345406 BX345406
   15   679.4     23.6   1040  13  BX345407                BX345407 BX345407
   16   664.2     23.1   1201  13  BX422753                BX422753 BX422753
   17     664     23.0    889  13  BQ691915                BQ691915 AGENCOURT
   18   653.6     22.7    934  10  BF311804                BF311804 601897316
   19   648.6     22.5   2532  29  AY411748                AY411748 Pan trogl
   20   646.4     22.4    859  10  BF311896                BF311896 601897733
   21   644.2     22.4    874  13  BQ689148                BQ689148 AGENCOURT
   22   604.4     21.0    756  13  BU612387                BU612387 UI-M-EW0-
   23   596.4     20.7    662   9  AL516580                AL516580 AL516580
   24     579     20.1   2775  29  AY401471                AY401471 Mus muscu
   25   573.4     19.9    604  10  BE314370                BE314370 601147261
   26   573.4     19.9   2775  29  AY401469                AY401469 Homo sapi
   27   554.6     19.3   1175  10  BF530640                BF530640 602071931
   28     548     19.0    788  14  CA317532                CA317532 UI-M-FW0-
   29   539.4     18.7    678  14  CA749784                CA749784 UI-M-FD0-
   30   524.4     18.2    824  12  BI737024                BI737024 603360874
   31     493     17.1   1030  12  BM479478                BM479478 AGENCOURT
   32   471.2     16.4    540  12  BI467799                BI467799 389583 MA
c  33   467.8     16.2    917  13  BX453456                BX453456 BX453456
c  34   463.6     16.1   1028  13  BX331575                BX331575 BX331575
   35   449.8     15.6    987  13  BX452510                BX452510 BX452510
   36     445     15.4   2507  29  AY401470                AY401470 Pan trogl
   37     433     15.0    471  13  BX282095                BX282095 BX282095
   38   430.6     14.9    675  14  CA315487                CA315487 UI-M-FW0-
   39   424.4     14.7    843  12  BG913440                BG913440 602811321
   40     424     14.7    780  12  BG470852                BG470852 602511705
   41     416     14.4    718  14  CA318180                CA318180 UI-M-FW0-
   42   415.4     14.4   1147  13  BU840446                BU840446 AGENCOURT
   43     413     14.3    960  12  BG329812                BG329812 602429458
   44   411.6     14.3   1111  12  BG298307                BG298307 602397080
   45   411.4     14.3    749  14  CF735417                CF735417 UI-M-HB0-

                              ALIGNMENTS

RESULT 1
AY406491
LOCUS       AY406491                2802 bp    DNA     linear   GSS 15-DEC-2003
DEFINITION  Homo sapiens UNC5C gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY406491
VERSION     AY406491.1  GI:39762465
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2802)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 2802)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
```

```
JOURNAL     Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1..2802
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     gene            <1..>2802
                     /gene="UNC5C"
                     /locus_tag="HCM2575"
ORIGIN

  Query Match             31.5%;  Score 907.8;  DB 29;  Length 2802;
  Best Local Similarity   60.3%;  Pred. No. 3.9e-142;
  Matches 1660;  Conservative    0;  Mismatches 1014;  Indels   78;  Gaps    7;

Qy         98 GCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGG 157
Db         66 GCTCGTGCTACCTGCCCTGGCCCTGCTCAGCGCCAGCGGCACTGGCTCCGCCGCCCAAGA 125

Qy        158 TGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCC 217
Db        126 TGATGACTTTTTTCATGAACTCCCAGAAACTTTTCCTTCTGATCCACCTGAGCCTCTGCC 185

Qy        218 CCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGT 277
Db        186 ACATTTCCTTATTGAGCCTGAAGAAGCTTATATTGTGAAGAATAAGCCTGTGAACCTGTA 245

Qy        278 GTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCA 337
Db        246 CTGTAAAGCAAGCCCTGCCACCCAGATCTATTTCAAGTGTAATAGTGAATGGGTTCATCA 305

Qy        338 GGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCGACCATGGAGGT 397
Db        306 GAAGGACCACATAGTAGATGAAAGAGTAGATGAAACTTCCGGTCTCATTGTCCGGGAAGT 365

Qy        398 CCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTG 457
Db        366 GAGCATTGAGATTTCGCGCCAGCAAGTGGAAGAACTCTTTGGACCTGAAGATTACTGGTG 425

Qy        458 CCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCAT 517
Db        426 CCAGTGTGTGGCCTGGAGCTCCGCGGGTACCACAAAGAGCCGGAAGGCGTATGTGCGCAT 485

Qy        518 AGC------CAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCT 571
Db        486 TGCATNNNNNNNNNNCGGAAGACATTTGAGCAGGAACCCCTAGGAAAGGAAGTGTCTTT 545

Qy        572 GGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGA 631
Db        546 GGAACAGGAAGTCTTACTCCAGTGTCGACCACCTGAAGGGATCCCAGTGGCTGAGGTGGA 605

Qy        632 GTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCG 691
Db        606 ATGGTTGAAAAATGAAGACATAATTGATCCCGTTGAAGATCGGAATTTTTATATTACTAT 665

Qy        692 GGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGT 751
Db        666 TGATCACAACCTCATCATAAAGCAGGCCCGACTCTCTGATACTGCAAATTACACCTGTGT 725

Qy        752 GGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTACGTGAA 811
Db        726 TGCCAAAAACATTGTTGCCAAGAGGAAAAGTACAACTGCCACTGTCATAGTCTATGTCAA 785

Qy        812 CGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTG 871
Db        786 CGGTGGCTGGTCCACCTGGACGGAGTGGTCTGTGTGTAACAGCCGCTGTGGACGAGGGTA 845

Qy        872 GCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGA 931
Db        846 TCAGAAACGTACAAGGACTTGTACCAACCCGGCACCACTCAATGGGGGTGCCTTCTGTGA 905

Qy        932 GGGGCAGAATGTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGCAGCTG 991
Db        906 AGGGCAGAGTGTGCAGAA---AATAGCCTGTACTACGTTATGCCCAGTGGATGGCAGGTG 962

Qy        992 GAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGA 1051
Db        963 GACGCCATGGAGCAAGTGGTCTACTTGTGGAACTGAGTGCACCCACTGGCGCAGGAGGGA 1022

Qy       1052 GTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACAC 1111
Db       1023 GTGCACGGCGCCAGCCCCCAAGAATGGAGGCAAGGACTGCGACGGCCTCGTCTTGCAATC 1082

Qy       1112 CCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCT 1171
Db       1083 CAAGAACTGCACTGATGGGCTTTGCATGCAGACTGCTCCTGATTCAGATGATGTTGCTCT 1142

Qy       1172 CTATGTGGG---CCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCAT 1228
Db       1143 CTATGTTGGGATTGTGATAGCAGTGATCGTTTGCCTGGCGATCTCTGTAGTTGTGGCCTT 1202

Qy       1229 CCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCT 1288
Db       1203 GTTTGTGTATCGGAAGAATCATCGTGACTTTGAGTCAGATATTATTGACTCTTCGGCACT 1262

Qy       1289 CACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCT 1348
Db       1263 CAATGGGGGCTTTCAGCCTGTGAACATCAAG---------GCAGCAAGACAAGATCTGCT 1313

Qy       1349 CACCATCCAGCCGGACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCA 1408
Db       1314 GGCTGTACCCCCAGACCTCACGTCAGCTGCAGCCATGTACAGAGGACCTGTCTATGCCCT 1373

Qy       1409 GGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGG 1468
Db       1374 GCATGACGTCTCAGACAAAATCCCAATGACCAACTCTCCAATTCTGGATCCACTGCCCAA 1433

Qy       1469 CGGCCGCCACACACTGCACCACAGCT------------CTCCCACCTCTGAGGCCGAGGA 1516
Db       1434 CCTGAAAATCAAAGTGTACAACACCTCAGGTGCTGTCACCCCCCAAGATGACCTCTCTGA 1493

Qy       1517 GTTCGTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAA 1576
Db       1494 GTTTACGTCCAAGCTGTCCCCTCAGATGACCCAGTCGTTGTTGGAGAATGAAGCCCTCAG 1553

Qy       1577 CATGA-----------------------------------------CCTATGGGACCTT 1594
Db       1554 CCTGAAGAACCAGAGTCTAGCAAGGCAGACTGATCCATCCTGTACCGCATTTGGCAGCTT 1613

Qy       1595 CAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCC 1654
Db       1614 CAACTCGCTGGGAGGTCACCTTATTGTTCCCAATTCAGGAGTCAGCTTGCTGATTCCCGC 1673

Qy       1655 AGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGA 1714
Db       1674 TGGGGCCATTCCCCAAGGGAGAGTCTACGAAATGTATGTGACTGTACACAGGAAAGAAAC 1733

Qy       1715 CGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACC 1774
Db       1734 TATGAGGCCACCCATGGATGACTCTCAGACACTTTTGACCCCTGTGGTGAGCTGTGGGCC 1793

Qy       1775 CCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAG 1834
Db       1794 CCCAGGAGCTCTGCTCACCCGCCCAGTCGTCCTCACTATGCATCACTGCGCAGACCCCAA 1853

Qy       1835 CCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGCAGGA 1894
Db       1854 TACCGAGGACTGGAAAATACTGCTCAAGAACCAGGCAGCACAGGGACAGTGGG---AGGA 1910

Qy       1895 TGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAG 1954
Db       1911 TGTGGTGGTGGTCGGGGAGGAAAACTTCACCACCCCCTGCTACATTCAGCTGGATGCAGA 1970
```

```
Qy       1955 TGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAG 2014
Db       1971 GGCCTGCCACATCCTCACAGAGAACCTCAGCACCTACGCCCTGGTAGGACATTCCACCAC 2030

Qy       2015 CGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCT 2074
Db       2031 CAAAGCGGCTGCGAAGCGCCTCAAGCTGGCCATCTTTGGGCCCCTGTGCTGCTCCTCGCT 2090

Qy       2075 CGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGT 2134
Db       2091 GGAGTACAGCATCCGAGTCTACTGTCTGGATGACACCCAGGATGCCCTGAAGGAAATTTT 2150

Qy       2135 GCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAA 2194
Db       2151 ACATCTTGAGAGACAGATGGGAGGACAGCTCCTAGAAGAACCTAAGGCTCTTCATTTTAA 2210

Qy       2195 GGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAG 2254
Db       2211 AGGCAGCACCCACAACCTGCGCCTGTCAATTCACGATATCGCCCATTCCCTCTGGAAGAG 2270

Qy       2255 TAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCG 2314
Db       2271 CAAATTGCTGGCTAAATATCAGGAAATTCCATTTTACCATGTTTGGAGTGGATCTCAAAG 2330

Qy       2315 GTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTG 2374
Db       2331 AAACCTGCACTGCACCTTCACTCTGGAAAGATTTAGCCTGAACACAGTGGAGCTGGTTTG 2390

Qy       2375 CAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACAT 2434
Db       2391 CAAACTCTGTGTGCGGCAGGTGGAAGGAGAAGGGCAGATCTTCCAGCTCAACTGCACCGT 2450

Qy       2435 CACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGC 2494
Db       2451 GTCAGAGGAACCTACTGGCATCGATTTGCCGCTGCTGGATCCTGCGAACACCATCACCAC 2510

Qy       2495 CCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAG 2554
Db       2511 GGTCACGGGGCCCAGTGCTTTCAGCATCCCTCTCCCTATCCGGCAGAAGCTCTGTAGCAG 2570

Qy       2555 CCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCT 2614
Db       2571 CCTGGATGCCCCCCAGACGAGAGGCCATGACTGGAGGATGCTGGCCCATAAGCTGAACCT 2630

Qy       2615 GGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCT 2674
Db       2631 GGACAGGTACTTGAATTACTTTGCCACCAAATCCAGCCCAACTGGCGTAATCCTGGATCT 2690

Qy       2675 GTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGG 2734
Db       2691 TTGGGAAGCACAGAACTTCCCAGATGGAAACCTGAGCATGCTGGCAGCTGTCTTGGAAGA 2750

Qy       2735 ACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2786
Db       2751 AATGGGAAGACATGAAACGGTGGTGTCCTTAGCAGCAGAAGGGCAGTATTAA 2802
```

```
RESULT 2
AY406493
LOCUS       AY406493                2791 bp    DNA     linear   GSS 15-DEC-2003
DEFINITION  Mus musculus UNC5C gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY406493
VERSION     AY406493.1  GI:39762467
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 2791)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 2791)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1..2791
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
     gene            <1..>2791
                     /gene="UNC5C"
                     /locus_tag="HCM2575"
ORIGIN

  Query Match             31.2%;  Score 898.6;  DB 29;  Length 2791;
  Best Local Similarity   61.5%;  Pred. No. 1.4e-140;
  Matches 1635;  Conservative    0;  Mismatches  939;  Indels   83;  Gaps    9;
```

```
Qy        190 TGCCTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACA 249
Db        158 TTCCATCTGACCCACCTGAGCCATTGCCACACTTCCTCATTGAGCCCGAGGAAGCTTACA 217

Qy        250 TCGTCAAGAACAAGCCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCT 309
Db        218 TTGTGAAGAACAAGCCTGTGAACCTGTATTGTAAAGCCAGCCCTGCCACCCAGATCTACT 277

Qy        310 TCAAGTGCAACGGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACG 369
Db        278 TCAAGTGCAACAGCGAGTGGGTTCATCAGAAGGACCACGTAGTAGACGAGAGAGTAGATG 337

Qy        370 GGAGCAGTGGTGAGCCGACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGA 429
Db        338 AAACCTCTGGTCTAATTGTGAGAGAAGTGAGCATTGAGATTTCACGCCAGCAGGTGGAGG 397

Qy        430 AGGTGTTCGGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCA 489
Db        398 AACTGTTTGGGCCTGAAGATTACTGGTGCCAGTGTGTGGCCTGGAGCTCAGCAGGCACTA 457

Qy        490 CCAAGAGTCAGAAGGCCTACATCCGCATAGCCAG------ATTGCGCAAGAACTTCGAGC 543
Db        458 CGAAGAGTCGGAAGGCATACGTGCGCATTGCGTTTCCAGATCTGCGGAAGACATTCGAGC 517

Qy        544 AGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCAC 603
Db        518 AGGAACCCTTGGGAAAGGAAGTGTCCTTGGAGCAGGAAGTCTTACTCCAGTGTCGGCCAC 577

Qy        604 CGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGT 663
Db        578 CTGAAGGGATCC-----------GTAGAATGGCTAAAGAATGAAGACATAATTGATCCTG 626

Qy        664 CCCTGGACCCCAATGTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCC 723
Db        627 CTGAAGATCGGAACTTTTATATTACTATCGATCACAACCTGATCATCAAGCAAGCCCGAC 686

Qy        724 TTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCG 783
Db        687 TCTCAGATACAGCAAATTATACCTGTGTTGCCAAAAATATTGTTGCCAAGAGAAAAAGCA 746

Qy        784 CCTCCGCTGCTGTCATCGTCTACGTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCG 843
Db        747 CCACAGCCACTGTCATCGTGTATGTTAATGGTGGCTGGTCCACCTGGACAGAGTGGTCTG 806
```

```
Qy    844 TCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGG 903
Db    807 TGTGTAACAGCCGCTGTGGGCGAGGATATCAGAAACGCACAAGAACCTGCACCAACCCAG 866
Qy    904 CGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCCT 963
Db    867 CCCCACTCAATGGTGGGGCCTTCTGTGAGGGGCAGAGTGTGCAGAA---AATAGCATGCA 923
Qy    964 CTCTGCTTGTCTCTGTGGACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGC 1023
Db    924 CTACGTTATGTCCAGTGGATGGTAGGTGGACTTCATGGAGCAAATGGTCAACCTGTGGGA 983
Qy   1024 TGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGG 1083
Db    984 CTGAATGCACCCACTGGCGCAGGAGGGAGTGTACAGCACCAGCCCCCAAGAACGGGGGTA 1043
Qy   1084 AGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGTACCAGTGACCTCTGTGTACACA 1143
Db   1044 AGGACTGTGATGGCCTGGTCCTCCAATCCAAGAACTGCACTGATGGGCTGTGCATGCAGG 1103
Qy   1144 GTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGG---CCTCATCGCCGTGGCCGTCT 1200
Db   1104 CTGCTCCTGACTCAGATGATGTGGCTCTCTACGTGGGGATTGTGATCGCTGTAACAGTCT 1163
Qy   1201 GCCTGGTCCTGCTGCTGCTTGTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGG 1260
Db   1164 GTCTGGCGATCACTGTTGTGGTGGCCCTGTTTGTGTATCGGAAGAACCACCGTGACTTTG 1223
Qy   1261 ACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGC 1320
Db   1224 AGTCTGACATCATTGACTCCTCAGCACTCAATGGCGGCTTTCAGCCTGTGAACATCAAGG 1283
Qy   1321 CCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCA 1380
Db   1284 CTGCCAGACAAGATCTCCTGGCTGTCCCCCCTGACC------TCACCTCAGCTGCAGCCA 1337
Qy   1381 CCTACCAGGGCAGTCTCTGTCCCCGGCAGGATG-----GGCCCAGCCCCAAGTTCCAGCT 1435
Db   1338 TGTACAGGGGACCTGTCTATGCTCTGCATGATGTCTCAGACAAAATCCCAATGACCAACT 1397
Qy   1436 CACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCACACACTGCACCACAGCTC 1495
Db   1398 CTCCAATTCTGGACCCACTACCCAACTTGAAAATCAAAGTGTACAACAGCTCAGGTGCTG 1457
Qy   1496 TCCCACCTC----TGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACCC------------ 1539
Db   1458 TCACTCCTCAGGATGACCTTGCCGAGTTCTCATCCAAACTGTCACCCCAGATGACCCAGT 1517
Qy   1540 ------------------------------AGAACTACTTCCGCTCCCTGCCCCGAGGCA 1569
Db   1518 CCTTGCTAGAGAATGAGGCCCTTAACCTGAAGAACCAGAGCCTCGCAAGACAGACTGACC 1577
Qy   1570 CCAGCAACATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATA 1629
Db   1578 CATCCTGCACAGCATTTGGTACCTTCAACTCTCTTGGGGGTCACCTCATCATTCCTAATT 1637
Qy   1630 CAGGTATCAGCCTCCTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCT 1689
Db   1638 CAGGAGTAAGCTTGCTGATTCCCGCTGGGGCCATTCCTCAGGGGAGAGTCTATGAAATGT 1697
Qy   1690 ACCTCACGCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGC 1749
Db   1698 ATGTGACTGTACACAGGAAAGAAAATATGAGGCCCCCCATGGAAGACTCTCAGACCCTAC 1757
Qy   1750 TGAGTCCCATCGTTAGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGG 1809
Db   1758 TTACCCCTGTGGTGAGCTGTGGGCCTCCTGGAGCTCTGCTGACCCGCCCTGTCATCCTCA 1817
Qy   1810 CTATGGACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGT 1869
Db   1818 CTCTGCATCACTGTGCAGACCCCAGCACCGAGGACTGGAAGATCCAGCTCAAAAACCAGG 1877
Qy   1870 CGTGCGAGGGCAGCTGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACC 1929
Db   1878 CAGTGCAGGGACAATGGAGGATGTTGTGGTG---GTTGGGGAGGAGAACTTCACAACCC 1934
Qy   1930 TCTACTACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCT 1989
Db   1935 CCTGTTACATTCAGCTGGATGCAGAGGCTTGCCATATCCTCACAGAGAACCTCAGTACCT 1994
Qy   1990 TTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGT 2049
Db   1995 ATGCCCTGGTTGGGCAGTCCACCACCAAAGCAGCTGCCAAGCGTCTTAAACTGGCCATCT 2054
Qy   2050 TTGCGCCGGTGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACA 2109
Db   2055 TTGGGCCCCTCTGCTGCTCTTCCCTGGAGTACAGCATTAGAGTCTACTGCCTGGATGACA 2114
Qy   2110 CCCACGATGCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCC 2169
Db   2115 CACAGGATGCCCTGAAGGAAGTTCTACAACTGGAGAGGCAAATGGGAGGACAGCTCCTAG 2174
Qy   2170 AGGAGCCACGGGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACG 2229
Db   2175 AAGAACCCAAGGCTCTTCATTTTAAAGGCAGCATCCACAACCTGCGCCTGTCTATTCATG 2234
Qy   2230 ATGTGCCCAGCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTT 2289
Db   2235 ACATCGCCCATTCCCTCTGGAAGAGCAAATTGCTGGCTAAGTATCAGGAAATTCCATTTT 2294
Qy   2290 ATCACATCTGGAATGGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCA 2349
Db   2295 ACCACATCTGGAGTGGCTCTCAAAGAAACCTCCACTGCACCTTCACTCTGGAAAGACTCA 2354
Qy   2350 GCCCCAGCACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGC 2409
Db   2355 GCCTAAACACAGTGGAACTGGTTTGCAAACTCTGTGTGCGGCAGGTTGAAGGAGAAGGGC 2414
Qy   2410 AGAGCTTCAGCATCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTC 2469
Db   2415 AGATCTTCCAGCTCAACTGTACTGTGTCAGAGGAACCTACTGGCATCGACTTACCTCTCC 2474
Qy   2470 TGGAGAGTGAAGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCC 2529
Db   2475 TGGACCCTGCTAGTACCATCACCACTGTCACCGGACCAAGTGCTTTCAGCATTCCTCTCC 2534
Qy   2530 TCATTCGGCAGAAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGC 2589
Db   2535 CTATCCGGCAGAAGCTATGCAGCAGCCTGGATGCCCCTCAAACAAGAGGCCATGACTGGA 2594
Qy   2590 GGACTCTGGCCCAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCA 2649
Db   2595 GGATGCTGGCCCATAAACTCAACCTGGACAGGTACTTGAATTACTTTGCCACCAAATCGA 2654
Qy   2650 GCCCCACAGCCATGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCA 2709
Db   2655 GCCCAACTGGCGTAATCCTGGATCTTTGGGAAGCACAGAACTTCCCAGATGGAAACCTGA 2714
Qy   2710 GCCAGCTGGCTGCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGT 2769
Db   2715 GCATGCTGGCAGCCGTCCTGGAAGAAATGGGAAGACATGAGACAGTGGTGTCCTTGGCAG 2774
Qy   2770 CGGAGGCTGAGTGCTGA 2786
Db   2775 CAGAAGGACAGTATTGA 2791
```

```
RESULT 3
AK031655
LOCUS       AK031655               3790 bp    mRNA    linear   HTC 18-SEP-2003
DEFINITION  Mus musculus 13 days embryo male testis cDNA, RIKEN full-length
            enriched library, clone:6030473H24 product:unc5 homolog (C.
            elegans) 3, full insert sequence.
ACCESSION   AK031655
VERSION     AK031655.1  GI:26327502
KEYWORDS    HTC; CAP trapper.
```

```
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4
  AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the
            FANTOM Consortium.
  TITLE     Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409, 685-690 (2001)
REFERENCE   5
  AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
            Group Phase I & II Team.
  TITLE     Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
REFERENCE   6  (bases 1 to 3790)
  AUTHORS   Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
            Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
            Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
            Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
            Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
            Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
            Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
            Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
            Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
            Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
            Muramatsu,M. and Hayashizaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT     cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
            Please visit our web site for further details.
            URL:http://genome.gsc.riken.go.jp/
            URL:http://fantom.gsc.riken.go.jp/.
FEATURES             Location/Qualifiers
     source          1..3790
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:6030473H24"
                     /db_xref="MGI:2394911"
                     /db_xref="taxon:10090"
                     /clone="6030473H24"
                     /sex="male"
                     /tissue_type="testis"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="13 days embryo"
     CDS             118..2970
                     /note="unnamed protein product; putative
                     unc5 homolog (C. elegans) 3 (MGD|MGI:1095412,
                     GB|NM_009472, evidence: BLASTN, 99%, match=464)"
                     /codon_start=1
                     /protein_id="BAC27495.1"
                     /db_xref="GI:26327503"
                     /translation="MRKGLRATAARCGLGIGYLLQMLVLPALALLSASGTGSAAQDDE
                     FFHELPETFPSDPPEPLPHFLIEPEEAYIVKNKPVNLYCKASPATQIYFKCNSEWVHQ
                     KDHVVDERVDETSGLIVREVSIEISRQQVEELFGPEDYWCQCVAWSSAGTTKSRKAYV
                     RIAYLRKTFEQEPLGKEVSLEQEVLLQCRPPEGIPVAEVEWLKNEDIIDPAEDRNFYI
                     TIDHNLIIKQARLSDTANYTCVAKNIVAKRKSTTATVIVYVNGGWSTWTEWSVCNSRC
                     GRGYQKRTRTCTNPAPLNGGAFCEGQSVQKIACTTLCPVDGRWTSWSKWSTCGTECTH
                     WRRRECTAPAPKNGGKDCDGLVLQSKNCTDGLCMQGFIYPISTEHRPQNEYGFSSAPD
                     SDDVALYVGIVIAVTVCLAITVVVALFVYRKNHRDFESDIIDSSALNGGFQPVNIKAA
                     RQDLLAVPPDLTSAAAMYRGPVYALHDVSDKIPMTNSPILDPLPNLKIKVYNSSGAVT
                     PQDDLAEFSSKLSPQMTQSLLENEALNLKNQSLARQTDPSCTAFGTFNSLGGHLIIPN
                     SGVSLLIPAGAIPQGRVYEMYVTVHRKENMRPPMEDSQTLLTPVVSCGPPGALLTRPV
                     ILTLHHCADPSTEDWKIQLKNQAVQGQWEDVVVVGEENFTTPCYIQLDAEACHILTEN
                     LSTYALVGQSTTKAAAKRLKLAIFGPLCCSSLEYSIRVYCLDDTQDALKEVLQLERQM
                     GGQLLEEPKALRFKGSIHNLRLSIHDIAHSLWKSKLLAKYQEIPFYHIWSGSQRNLHC
                     TFTLERLSLNTVELVCKLCVRQVEGEGQIFQLNCTVSEEPTGIDLPLLDPASTITTVT
                     GPSAFSIPLPIRQKLCSSLDAPQTRGHDWRMLAHKLNLDRYLNYFATKSSPTGVILDL
                     WEAQNFPDGNLSMLAAVLEEMGRHETVVYLAAEGQY"
ORIGIN

  Query Match             30.3%;  Score 871.6;  DB 11;  Length 3790;
  Best Local Similarity   60.6%;  Pred. No. 5.5e-136;
  Matches 1641;  Conservative    0;  Mismatches  944;  Indels  123;  Gaps    8;

Qy        190 TGCCTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACA 249

Db        275 TTCCATCTGACCCACCTGAGCCATTGCCACACTTCCTCATTGAGCCCGAGGAAGCTTACA 334

Qy        250 TCGTCAAGAACAAGCCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCT 309

Db        335 TTGTGAAGAACAAGCCTGTGAACCTGTATTGTAAAGCCAGCCCTGCCACCCAGATCTACT 394

Qy        310 TCAAGTGCAACGGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACG 369

Db        395 TCAAGTGCAACAGCGAGTGGGTTCATCAGAAGGACCACGTAGTAGACGAGAGAGTAGATG 454

Qy        370 GGAGCAGTGGTGAGCCGACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGA 429

Db        455 AAACCTCTGGTCTAATTGTGAGAGAAGTGAGCATTGAGATTTCACGCCAGCAGGTGGAGG 514

Qy        430 AGGTGTTCGGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCA 489

Db        515 AACTGTTTGGGCCTGAAGATTACTGGTGCCAGTGTGTGGCCTGGAGCTCAGCAGGCACTA 574

Qy        490 CCAAGAGTCAGAAGGCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGC 549

Db        575 CGAAGAGTCGGAAGGCATACGTGCGCATTGCGTATCTGCGGAAGACATTCGAGCAGGAAC 634

Qy        550 CGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGG 609

Db        635 CCTTGGGAAAGGAAGTGTCCTTGGAGCAGGAAGTCTTACTCCAGTGTCGGCCACCTGAAG 694

Qy        610 GCATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGG 669

Db        695 GGATCCCAGTGGCTGAGGTAGAATGGCTAAAGAATGAAGACATAATTGATCCTGCTGAAG 754
```

```
Qy    670 ACCCCAATGTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTG 729
Db    755 ATCGGAACTTTTATATTACTATCGATCACAACCTGATCATCAAGCAAGCCCGACTCTCAG 814

Qy    730 ACACGGCCAACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCG 789
Db    815 ATACAGCAAATTATACCTGTGTTGCCAAAAATATTGTTGCCAAGAGAAAAAGCACCACAG 874

Qy    790 CTGCTGTCATCGTCTACGTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCA 849
Db    875 CCACTGTCATCGTGTATGTTAATGGTGGCTGGTCCACCTGGACAGAGTGGTCTGTGTGTA 934

Qy    850 GCGCCAGCTGTGGGCGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTC 909
Db    935 ACAGCCGCTGTGGGCGAGGATATCAGAAACGCACAAGAACCTGCACCAACCCAGCCCCAC 994

Qy    910 TCAACGGGGGCGCTTTCTGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCCTCTCTGC 969
Db    995 TCAATGGTGGGGCCTTCTGTGAGGGGCAGAGTGTGCAGAA---AATAGCATGCACTACGT 1051

Qy    970 TTGTCTCTGTGGACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACT 1029
Db   1052 TATGTCCAGTGGATGGTAGGTGGACTTCATGGAGCAAATGGTCAACCTGTGGGACTGAAT 1111

Qy   1030 GCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGT 1089
Db   1112 GCACCCACTGGCGCAGGAGGGAGTGTACAGCACCAGCCCCCAAGAACGGGGGTAAGGACT 1171

Qy   1090 GCCAGGGCACTGACCTGGACACCCGCAACTGTACCAGTGACCTCTGTGTACACAG----- 1144
Db   1172 GTGATGGCCTGGTCCTCCAATCCAAGAACTGCACTGATGGGCTGTGCATGCAGGGATTCA 1231

Qy   1145 ----------------------------------------------------TGCTTCTG 1152
Db   1232 TTTACCCCATTTCAACTGAGCACAGACCCCAGAATGAATATGGATTTTCTTCTGCTCCTG 1291

Qy   1153 GCCCTGAGGACGTGGCCCTCTATGTGGG---CCTCATCGCCGTGGCCGTCTGCCTGGTCC 1209
Db   1292 ACTCAGATGATGTGGCTCTCTACGTGGGGATTGTGATCGCTGTAACAGTCTGTCTGGCGA 1351

Qy   1210 TGCTGCTGCTTGTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATG 1269
Db   1352 TCACTGTTGTGGTGGCCCTGTTTGTGTATCGGAAGAACCACCGTGACTTTGAGTCTGACA 1411

Qy   1270 TGGCTGACTCGTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAG 1329
Db   1412 TCATTGACTCCTCAGCACTCAATGGCGGCTTTCAGCCTGTGAACATCAAGGCTGCCAGAC 1471

Qy   1330 CAGACAACCCCCATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCACCTACCAGG 1389
Db   1472 AAGATCTCCTGGCTGTCCCCCCTGACC------TCACCTCAGCTGCAGCCATGTACAGGG 1525

Qy   1390 GCAGTCTCTGTCCCCGGCAGGATG-----GGCCCAGCCCCAAGTTCCAGCTCACCAATGG 1444
Db   1526 GACCTGTCTATGCTCTGCATGATGTCTCAGACAAAATCCCAATGACCAACTCTCCAATTC 1585

Qy   1445 GCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTC 1504
Db   1586 TGGACCCACTACCCAACTTGAAAATCAAAGTGTACAACAGCTCAGGTGCTGTCACTCCTC 1645

Qy   1505 ----TGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACCC--------------------- 1539
Db   1646 AGGATGACCTTGCCGAGTTCTCATCCAAACTGTCACCCCAGATGACCCAGTCCTTGCTAG 1705

Qy   1540 ---------------------AGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACA 1578
Db   1706 AGAATGAGGCCCTTAACCTGAAGAACCAGAGCCTCGCAAGACAGACTGACCCATCCTGCA 1765

Qy   1579 TGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCA 1638
Db   1766 CAGCATTTGGTACCTTCAACTCTCTTGGGGTCACCTCATCATTCCTAATTCAGGAGTAA 1825

Qy   1639 GCCTCCTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGC 1698
Db   1826 GCTTGCTGATTCCCGCTGGGGCCATTCCTCAGGGGAGAGTCTATGAAATGTATGTGACTG 1885

Qy   1699 TGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCA 1758
Db   1886 TACACAGGAAAGAAAATATGAGGCCCCCCATGGAAGACTCTCAGACCCTACTTACCCCTG 1945

Qy   1759 TCGTTAGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACC 1818
Db   1946 TGGTGAGCTGTGGGCCTCCTGGAGCTCTGCTGACCCGCCCTGTCATCCTCACTCTGCATC 2005

Qy   1819 ACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGG 1878
Db   2006 ACTGTGCAGACCCCAGCACCGAGGACTGGAAGATCCAGCTCAAAAACCAGGCAGTGCAGG 2065

Qy   1879 GCAGCTGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACT 1938
Db   2066 GACAATGGGAGGATGTTGTGGTG---GTTGGGGAGGAGAACTTCACAACCCCCTGTTACA 2122

Qy   1939 GCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGG 1998
Db   2123 TTCAGCTGGATGCAGAGGCTTGCCATATCCTCACAGAGAACCTCAGTACCTATGCCCTGG 2182

Qy   1999 TGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGG 2058
Db   2183 TTGGGCAGTCCACCACCAAAGCAGCTGCCAAGCGTCTTAAACTGGCCATCTTTGGGCCCC 2242

Qy   2059 TGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATG 2118
Db   2243 TCTGCTGCTCTTCCCTGGAGTACAGCATTAGAGTCTACTGCCTGGATGACACACAGGATG 2302

Qy   2119 CACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCAC 2178
Db   2303 CCCTGAAGGAAGTTCTACAACTGGAGAGGCAAATGGGAGGACAGCTCCTAGAAGAACCCA 2362

Qy   2179 GGGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCA 2238
Db   2363 AGGCTCTTCGTTTTAAAGGCAGCATCCACAACCTGCGCCTGTCTATTCATGACATCGCCC 2422

Qy   2239 GCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCT 2298
Db   2423 ATTCCCTCTGGAAGAGCAAATTGCTGGCTAAGTATCAGGAAATTCCATTTTACCACATCT 2482

Qy   2299 GGAATGGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCA 2358
Db   2483 GGAGTGGCTCTCAAAGAAACCTCCACTGCACCTTCACTCTGGAAAGACTCAGCCTAAACA 2542

Qy   2359 CTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCA 2418
Db   2543 CAGTGGAACTGGTTTGCAAACTCTGTGTGCGGCAGGTTGAAGGAGAAGGGCAGATCTTCC 2602

Qy   2419 GCATCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTG 2478
Db   2603 AGCTCAACTGTACTGTGTCAGAGGAACCTACTGGCATCGACTTACCTCTCCTGGACCCTG 2662

Qy   2479 AAGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGC 2538
Db   2663 CTAGTACCATCACCACTGTCACCGGACCAAGTGCTTTCAGCATTCCTCTCCCTATCCGGC 2722

Qy   2539 AGAAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGG 2598
Db   2723 AGAAGCTATGCAGCAGCCTGGATGCCCCTCAAACAAGAGGCCATGACTGGAGGATGCTGG 2782

Qy   2599 CCCAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAG 2658
Db   2783 CCCATAAACTCAACCTGGACAGGTACTTGAATTACTTTGCCACCAAATCGAGCCCAACTG 2842

Qy   2659 CCATGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGG 2718
Db   2843 GCGTAATCCTGGATCTTTGGGAAGCACAGAACTTCCCAGATGGAAACCTGAGCATGCTGG 2902

Qy   2719 CTGCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTG 2778
```

```
Db        2903 CAGCCGTCCTGGAAGAAATGGGAAGACATGAGACAGTGGTGTACTTGGCAGCAGAAGGAC 2962

Qy        2779 AGTGCTGA 2786
               |||  |||
Db        2963 AGTATTGA 2970

RESULT 4
AK018177
LOCUS       AK018177                3866 bp    mRNA    linear   HTC 20-SEP-2003
DEFINITION  Mus musculus adult male medulla oblongata cDNA, RIKEN full-length
            enriched library, clone:6330415E02 product:TRANSMEMBRANE RECEPTOR
            UNC5H2 homolog [Rattus norvegicus], full insert sequence.
ACCESSION   AK018177
VERSION     AK018177.1  GI:12857775
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4
  AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the
            FANTOM Consortium.
  TITLE     Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409, 685-690 (2001)
REFERENCE   5
  AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
            Group Phase I & II Team.
  TITLE     Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
REFERENCE   6  (bases 1 to 3866)
  AUTHORS   Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H.,
            Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y.,
            Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K.,
            Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M.,
            Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,
            Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K.,
            Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C.,
            Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
            Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y.,
            Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T.,
            Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K.,
            Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT     Please visit our web site (http://genome.gsc.riken.go.jp/) for
            further details.
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues. First strand cDNA was primed with a primer
            [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was
            prepared by using trehalose thermo-activated reverse transcriptase
            and subsequently enriched for full-length by cap-trapper. cDNA went
            through one round of normalization to Rot = 10.0 and subtraction to
            Rot = 100.0. Second strand cDNA was prepared with the primer
            adapter of sequence [5'
            GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC 3']. cDNA was cleaved
            with BamHI and XhoI. Vector: a modified pBluescript KS(+) after
            bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3'
            end: BamHI. Host: DH10B.
FEATURES             Location/Qualifiers
     source          1..3866
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:6330415E02"
                     /db_xref="MGI:1910526"
                     /db_xref="taxon:10090"
                     /clone="6330415E02"
                     /sex="male"
                     /tissue_type="medulla oblongata"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="adult"
     CDS             417..3254
                     /note="unnamed protein product; TRANSMEMBRANE RECEPTOR
                     UNC5H2 homolog [Rattus norvegicus] (SPTR|O08722, evidence:
                     FASTY, 96.5%ID, 100%length, match=2835)
                     putative"
                     /codon_start=1
                     /protein_id="BAB31108.1"
                     /db_xref="GI:12857776"
                     /translation="MRARSGVRSALLLALLLCWDPTPSLAGVDSAGQVLPDSYPSAPA
                     EQLPYFLLEPQDAYIVKNKPVELHCRAFPATQIYFKCNGEWVSQNDHVTQESLDEATG
                     LRVREVQIEVSRQQVEELFGLEDYWCQCVAWSSSGTTKSRRAYIRIAYLRKNFDQEPL
                     AKEVPLDHEVLLQCRPPEGVPVAEVEWLKNEDVIDPAQDTNFLLTIDHNLIIRQARLS
                     DTANYTCVAKNIVAKRRSTAATVIVYVNGGWSSWAEWSPCSNRCGRGWQKRTRTCTNP
                     APLNGGAFCEGQAFQKTACTTVCPVDGAWTEWSKWSACSTECAHWRSRECMAPPPQNG
                     GRDCSGTLLDSKNCTDGLCVLNQRTLNDPKSHPLETSGDVALYAGLVVAVFVVVAVLM
                     AEGVIVYRRNCRDFDTDITDSSAALTGGFHPVNFKTARPNNPQLLHPSAPPDLTASAG
                     IYRGPVYALQDSADKIPMTNSPLLDPLPSLKIKVYNSSTIGSGSGLADGADLLGVLPP
                     GTYPGDFSRDTHFLHLRSASLGSQHLLGLPRDPSSSVSGTFGCLGGRLSLPGTGVSLL
                     VPNGAIPQGKFYDLYLHINKAESTLPLSEGSQTVLSPSVTCGPTGLLLCRPVVLTVPH
                     CAEVIAGDWIFQLKTQAHQGHWEEVVTLDEETLNTPCYCQLEAKSCHILLDQLGSYVF
                     MGESYSRSAVKRLQLAIFAPALCTSLEYSLRVYCLEDTPVALKEVLELERTLGGYLVE
                     EPKPLLFKDSYHNLRLSLHDIPHAHWRSKLLAKYQEIPFYHVWNGSQRALHCTFTLER
                     HSLASTEFTCKVCVRQVEGEGQIFQLHTTLAETPAGSLDALCSAPGNAITTQLGPYAF
                     KIPLSIRQKICSSLDAPDSRGNDWRLLAQKLSMDRYLNYFATKASPTGVILDLWEARQ
                     QDDGDLNSLASALEEMGKSEMLVAMATDGDC"
ORIGIN

  Query Match             28.4%;  Score 818.4;  DB 11;  Length 3866;
  Best Local Similarity   59.6%;  Pred. No. 4.9e-127;
  Matches 1623;  Conservative    0;  Mismatches  961;  Indels  138;  Gaps    9;

Qy         198 GCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAG 257
               |||    |||| | ||| ||  |||||||  ||||||  ||||  |   ||||||| |||
Db         540 GCCCCTGCGGAGCAGCTGCCGTACTTCCTATTGGAGCCACAGGACGCCTACATCGTAAAG 599

Qy         258 AACAAGCCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGC 317
               |||||||||||    ||      ||||  ||| | || ||||||||||||| ||||||||
```

```
Db    600 AACAAGCCAGTGGAACTGCACTGCAGAGCCTTCCCTGCCACGCAGATCTACTTCAAGTGT 659

Qy    318 AACGGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGT 377
Db    660 AATGGCGAGTGGGTCAGCCAGAATGACCACGTCACACAGGAGAGCCTGGATGAGGCCACA 719

Qy    378 GGTGAGCCGACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTC 437
Db    720 GGCTTGCGGGTGCGAGAGGTGCAGATCGAGGTGTCACGGCAGCAAGTGGAGGAACTCTTC 779

Qy    438 GGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGT 497
Db    780 GGGCTCGAGGACTACTGGTGCCAGTGCGTGGCCTGGAGCTCTTCGGGAACTACCAAGAGT 839

Qy    498 CAGAAGGCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCC 557
Db    840 CGCCGAGCCTACATCCGCATTGCCTACTTGCGCAAGAACTTTGACCAGGAGCCTCTGGCC 899

Qy    558 AAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCT 617
Db    900 AAGGAGGTACCCTTGGATCATGAGGTCCTTCTGCAGTGCCGCCCACCGGAGGGAGTGCCT 959

Qy    618 CCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAAT 677
Db    960 GTGGCTGAGGTGGAATGGCTCAAGAATGAAGATGTCATTGACCCCGCTCAGGACACTAAC 1019

Qy    678 GTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCC 737
Db   1020 TTCCTGCTCACCATTGACCACAACCTCATCATCCGCCAGGCGCGCCTCTCAGACACGGCC 1079

Qy    738 AACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTC 797
Db   1080 AACTACACCTGTGTGGCCAAGAATATCGTGGCCAAGCGCCGGAGCACCGCGGCCACAGTC 1139

Qy    798 ATCGTCTACGTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGC 857
Db   1140 ATCGTCTATGTGAATGGAGGCTGGTCCAGCTGGGCAGAGTGGTCACCCTGTTCCAATCGC 1199

Qy    858 TGTGGGCGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGG 917
Db   1200 TGTGGCCGAGGCTGGCAGAAGCGTACTCGGACCTGCACCAATCCAGCCCCACTCAATGGA 1259

Qy    918 GGCGCTTTCTGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCT 977
Db   1260 GGCGCCTTCTGTGAGGGACAGGCCTTCCAGAAGACAGCTTGCACCACCGTGTG---CCCA 1316

Qy    978 GTGGACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCAC 1037
Db   1317 GTGGATGGAGCGTGGACCGAGTGGAGCAAGTGGTCTGCCTGCAGCACAGAGTGTGCGCAC 1376

Qy   1038 TGGCGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGC 1097
Db   1377 TGGCGCAGCCGCGAGTGCATGGCACCGCCACCCCAGAACGGAGGCCGTGACTGCAGCGGG 1436

Qy   1098 ACTGACCTGGACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCC-- 1155
Db   1437 ACGCTACTTGACTCCAAGAACTGCACTGATGGGCTGTGCGTGCTGAATCAGAGAACTCTA 1496

Qy   1156 ------------------------------CTGAGGACGTGGCCCTCTATGTGGGCCTC 1184
Db   1497 AACGACCCTAAAAGCCACCCCCTGGAGACATCGGGAGATGTGGCACTGTACGCAGGCCTT 1556

Qy   1185 ATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTC---CTCATCCTCGTTTATTGC 1241
Db   1557 GTGGTGGCCGTCTTTGTGGTGGTAGCGGTTCTCATGGCCGAGGGAGTGATCGTATACCGG 1616

Qy   1242 CGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGT---CCATTCTCACCTCAGGC 1298
Db   1617 AGAAACTGCCGGGACTTCGACACGGACATCACCGACTCCTCTGCGGCCCTCACTGGTGGC 1676

Qy   1299 TTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCT------CACC 1352
Db   1677 TTCCACCCTGTCAACTTCAAGACTGCAAGGCCCAACAACCCGCAGCTCCTGCACCCGTCC 1736

Qy   1353 ATCCAGCCGGACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGAT 1412
Db   1737 GCCCCTCCAGACCTAACGGCCAGTGCTGGCATCTACCGCGGGCCTGTGTATGCCCTGCAG 1796

Qy   1413 GGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCT------------------- 1453
Db   1797 GACTCCGCCGACAAGATCCCCATGACTAATTCGCCCCTGCTGGATCCCCTGCCCAGCCTC 1856

Qy   1454 -----------------------------------------------------------C 1454
Db   1857 AAGATCAAGGTCTATAACTCCAGCACCATCGGTTCTGGGTCTGGCCTGGCTGATGGAGCC 1916

Qy   1455 AGCCCCCTGGGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAG 1514
Db   1917 GACCTGCTGGGTGTCCTCCCGCCGGGCACGTACCCAGGCGATTTCTCCCGGGACACCCAT 1976

Qy   1515 GAGTTCGTCTCCCGCCTCTCCACCCAG------AACTACTTCCGCTCCCTGCCCCGAGGC 1568
Db   1977 TTCCTGCACCTGCGCAGTGCCAGCCTTGGTTCCCAGCACCTCCTGGGCCTACCTCGGGAC 2036

Qy   1569 ACCAGCAACATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAAT 1628
Db   2037 CCCAGCAGCAGTGTCAGCGGCACCTTTGGTTGCCTGGGAGGAAGGCTGAGCCTCCCCGGC 2096

Qy   1629 ACAGGTATCAGCCTCCTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATC 1688
Db   2097 ACAGGGGTCAGCCTGTTGGTACCAAATGGAGCCATTCCCCAGGGCAAGTTCTATGACCTG 2156

Qy   1689 TACCTCACGCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTG 1748
Db   2157 TATCTACATATCAACAAGGCCGAAAGCACCCTCCCACTTTCAGAAGGTTCCCAGACAGTA 2216

Qy   1749 CTGAGTCCCATCGTTAGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTG 1808
Db   2217 TTGAGCCCCTCGGTGACCTGTGGGCCCACAGGCCTACTCCTGTGCCGCCCTGTCGTCCTC 2276

Qy   1809 GCTATGGACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAG 1868
Db   2277 ACCGTGCCCCACTGTGCTGAAGTCATCGCTGGAGACTGGATCTTTCAGCTCAAGACCCAG 2336

Qy   1869 TCGTGCGAGGGCAGCTGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCAC 1928
Db   2337 GCCCATCAGGGCCACTGGG---AGGAGGTGGTGACCTTGGATGAGGAGACCCTCAACACA 2393

Qy   1929 CTCTACTACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGC 1988
Db   2394 CCCTGCTACTGCCAGCTGGAGGCTAAGTCCTGCCACATCCTGCTGGACCAGCTGGGTTCC 2453

Qy   1989 TTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTG 2048
Db   2454 TACGTATTCATGGGCGAGTCCTACTCTCGCTCTGCAGTCAAGCGGCTCCAGCTGGCCATC 2513

Qy   2049 TTTGCGCCGGTGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGAC 2108
Db   2514 TTCGCCCCAGCCCTCTGCACCTCCCTGGAGTATAGCCTCAGGGTCTACTGTCTGGAGGAC 2573

Qy   2109 ACCCACGATGCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATC 2168
Db   2574 ACACCTGTAGCACTGAAGGAGGTCCTGGAGCTGGAGAGGACTCTGGGTGGCTACTTGGTG 2633

Qy   2169 CAGGAGCCACGGGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCAC 2228
Db   2634 GAGGAGCCCAAGCCTTTGCTCTTTAAGGACAGTTACCACAACCTACGCCTCTCCCTCCAT 2693

Qy   2229 GATGTGCCCAGCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTT 2288
Db   2694 GACATCCCCCATGCCCACTGGAGGAGCAAACTACTGGCCAAGTACCAGGAGATTCCCTTC 2753

Qy   2289 TATCACATCTGGAATGGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTC 2348
Db   2754 TACCACGTCTGGAATGGCAGCCAGAGAGCCCTGCACTGCACTTTCACCCTGGAGAGGCAT 2813
```

```
Qy       2349 AGCCCCAGCACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGG 2408
Db       2814 AGCCTGGCCTCCACGGAGTTCACCTGTAAGGTCTGCGTGCGGCAGGTCGAAGGGGAAGGC 2873

Qy       2409 CAGAGCTTCAGCATCAACTTCAACATCACC---AAGGACACAAGGTTTGCTGAGCTGCTG 2465
Db       2874 CAGATTTTCCAGCTGCACACAACGTTGGCCGAGACGCCTGCTGGCTCCCTGGATGCTCTC 2933

Qy       2466 GCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCC 2525
Db       2934 TGCTCTGCCCCGGGCAATGCCATCACCACCCAGCTGGGACCCTATGCCTTCAAGATACCC 2993

Qy       2526 TTCCTCATTCGGCAGAAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGAC 2585
Db       2994 CTGTCCATCCGCCAAAAGATCTGCAGCAGCCTGGACGCCCCCGACTCCCGGGGCAACGAC 3053

Qy       2586 TGGCGGACTCTGGCCCAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAG 2645
Db       3054 TGGAGGCTGTTGGCGCAGAAGCTGTCCATGGACCGGTACCTAAACTACTTCGCCACCAAA 3113

Qy       2646 CCCAGCCCCACAGCCATGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAAC 2705
Db       3114 GCTAGTCCCACAGGTGTCATCTTAGACCTCTGGGAAGCTCGGCAACAGGATGACGGGGAC 3173

Qy       2706 CTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACA 2765
Db       3174 CTCAACAGCCTGGCCAGTGCCTTGGAGGAGATGGGCAAGAGTGAGATGCTGGTAGCCATG 3233

Qy       2766 GTGTCGGAGGCTGAGTGCTGAG 2787
Db       3234 GCCACAGATGGCGATTGCTGAG 3255
```

RESULT 5
AY411747

```
LOCUS       AY411747                2532 bp    DNA     linear   GSS 12-DEC-2003
DEFINITION  Homo sapiens HCM4327 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY411747
VERSION     AY411747.1  GI:39767715
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2532)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 2532)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1..2532
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     gene            <1..>2532
                     /locus_tag="HCM4327"
ORIGIN
```

```
  Query Match             26.5%;  Score 763.2;  DB 29;  Length 2532;
  Best Local Similarity   60.0%;  Pred. No. 7.6e-118;
  Matches 1515;  Conservative    0;  Mismatches  873;  Indels  138;  Gaps    9;

Qy        393 GAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATAC 452
Db         13 GAGGTGCAGATCGAGGTGTCGCGGCAGCAGGTGGAGGAGCTCTTTGGGCTGGAGGATTAC 72

Qy        453 TGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATC 512
Db         73 TGGTGCCAGTGCGTGGCCTGGAGCTCCGCGGGCACCACCAAGAGTCGCCGAGCCTACGTC 132

Qy        513 CGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTG 572
Db        133 CGCATCGCCTACCTGCGCAAGAACTTCGATCAGGAGCCTCTGGGCAAGGAGGTGCCCCTG 192

Qy        573 GAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAG 632
Db        193 GACCATGAGGTTCTCCTGCAGTGCCGCCCGCCGGAGGGGGTGCCTGTGGCCGAGGTGGAA 252

Qy        633 TGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGG 692
Db        253 TGGCTCAAGAATGAGGATGTCATCGACCCCACCCAGGACACCAACTTCCTGCTCACCATC 312

Qy        693 GAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTG 752
Db        313 GACCACAACCTCATCATCCGCCAGGCCCGCCTGTCGGACACTGCCAACTATACCTGCGTG 372

Qy        753 GCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTACGTGAAC 812
Db        373 GCCAAGAACATCGTGGCCAAACGCCGGAGCACCACTGCCACCGTCATCGTCTACGTGAAT 432

Qy        813 GGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGG 872
Db        433 GGCGGCTGGTCCAGCTGGGCAGAGTGGTCACCCTGCTCCAACCGCTGTGGCCGAGGCTGG 492

Qy        873 CAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAG 932
Db        493 CAGAAGCGCACCCGGACCTGCACCAACCCCGCTCCACTCAACGGAGGGGCCTTCTGCGAG 552

Qy        933 GGGCAGAATGTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGCAGCTGG 992
Db        553 GGCCAGGCATTCCAGAA---GACCGCCTGCACCACCATCTGCCCAGTCGATGGGGCGTGG 609

Qy        993 AGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAG 1052
Db        610 ACGGAGTGGAGCAAGTGGTCAGCCTGCAGCACTGAGTGTGCCCACTGGCGTAGCCGCGAG 669

Qy       1053 TGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACC 1112
Db        670 TGCATGGCGCCCCCACCCCAGAACGGAGGCCGTGACTGCAGCGGGACGCTGCTCGACTCT 729

Qy       1113 CGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCT-------------------- 1151
Db        730 AAGAACTGCACAGATGGGCTGTGCATGCAAAATAAGAAAACTCTAAGCGACCCCAACAGC 789

Qy       1152 -----------GGCCCTGAGGACGTGGCCCTCTATGTGGGCCTC---ATCGCCGTGGCC 1196
Db        790 CACCTGCTGGAGGCCTCAGGGGATGCGGCGCTGTATGCGGGGCTCGTGGTGGCCATCTTC 849

Qy       1197 GTCTGCCTGGTCCTGCTGCTGCTTGTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGG 1256
Db        850 GTGGTCGTGGCAATCCTCATGGCGGTGGGGGTGGTGGTGTACCGCCGCAACTGCCGTGAC 909

Qy       1257 CTGGACTCAGATGTGGCTGACTCGTCCATT---CTCACCTCAGGCTTCCAGCCCGTCAGC 1313
Db        910 TTCGACACAGACATCACTGACTCATCTGCTGCCCTGACTGGTGGTTTCCACCCCGTCAAC 969

Qy       1314 ATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCT------CACCATCCAGCCGGACCTC 1367
Db        970 TTTAAGACGGCAAGGCCCAGCAACCCGCAGCTCCTACACCCCTCTGTGCCTCCTGACCTG 1029

Qy       1368 AGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCCAGCCCCAAG 1427
```

```
Db   1030 ACAGCCAGCGCCGGCATCTACCGCGGACCCGTGTATGCCCTGCAGGACTCCACCGACAAA 1089

Qy   1428 TTCCAGCTCACCAATGGCACCTGCTCAGCCCC---------------------------- 1460

Db   1090 ATCCCCATGACCAACTCTCCTCTGCTGGACCCCTTACCCAGCCTTAAGGTCAAGGTCTAC 1149

Qy   1461 ---------------------------------------------------CTGGGTGGC 1469

Db   1150 AGCTCCAGCACCACGGGCTCTGGGCCAGGCCTGGCAGATGGGGCTGACCTGCTGGGGGTC 1209

Qy   1470 GGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGC 1529

Db   1210 TTGCCGCCTGGCACATACCCTAGCGATTTCGCCCGGGACACCCACTTCCTGCACCTGCGC 1269

Qy   1530 CTCTCCACCCAGAACTAC------TTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC 1583

Db   1270 AGCGCCAGCCTCGGTTCCCAGCAGCTCTTGGGCCTGCCCCGAGACCCAGGGAGCAGCGTC 1329

Qy   1584 TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC 1643

Db   1330 AGCGGCACCTTTGGCTGCCTGGGTGGAGGCTCAGCATCCCCGGCACAGGGGTCAGCTTG 1389

Qy   1644 CTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC 1703

Db   1390 CTGGTGCCCAATGGAGCCATTCCCCAGGGCAAGTTCTACGAGATGTATCTACTCATCAAC 1449

Qy   1704 AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT 1763

Db   1450 AAGGCAGAAAGTACCCTCCCGCTTTCAGAAGGGACCCAGACAGTATTGAGCCCCTCGGTG 1509

Qy   1764 AGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT 1823

Db   1510 ACCTGTGGACCCACAGGCCTCCTGCTGTGCCGCCCCGTCATCCTCACCATGCCCCACTGT 1569

Qy   1824 GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC 1883

Db   1570 GCCGAAGTCAGTGCCCGTGACTGGATCTTTCAGCTCAAGACCCAGGCCCACCAGGGCCAC 1629

Qy   1884 TGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG 1943

Db   1630 TGGG---AGGAGGTGGTGACCCTGGATGAGGAGACCCTGAACACACCCTGCTACTGCCAG 1686

Qy   1944 CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA 2003

Db   1687 CTGGAGCCCAGGGCCTGTCACATCCTGCTGGACCAGCTGGGCACCTACGTGTTCACGGGC 1746

Qy   2004 GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC 2063

Db   1747 GAGTCCTATTCCCGCTCAGCAGTCAAGCGGCTCCAGCTGGCCGTCTTCGCCCCCGCCCTC 1806

Qy   2064 TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC 2123

Db   1807 TGCACCTCCCTGGAGTACAGCCTCCGGGTCTACTGCCTGGAGGACACGCCTGTAGCACTG 1866

Qy   2124 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTC 2183

Db   1867 AAGGAGGTGCTGGAGCTGGAGCGGACTCTGGGCGGATACTTGGTGGAGGAGCCGAAACCG 1926

Qy   2184 CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCC 2243

Db   1927 CTAATGTTCAAGGACAGTTACCACAACCTGCGCCTCTCCCTCCATGACCTCCCCCATGCC 1986

Qy   2244 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2303

Db   1987 CATTGGAGGAGCAAGCTGCTGGCCAAATACCAGGAGATCCCCTTCTATCACATTTGGAGT 2046

Qy   2304 GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2363

Db   2047 GGCAGCCAGAAGGCCCTCCACTGCACTTTCACCCTGGAGAGGCACAGCTTGGCCTCCACA 2106

Qy   2364 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2423

Db   2107 GAGCTCACCTGCAAGATCTGCGTGCGGCAAGTGGAAGGGGAGGGCCAGATATTCCAGCTG 2166

Qy   2424 AACTTCAACATCACCAAG---GACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAA 2480

Db   2167 CATACCACTCTGGCAGAGACACCTGCTGGCTCCCTGGACACTCTCTGCTCTGCCCCTGGC 2226

Qy   2481 GCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAG 2540

Db   2227 AGCACTGTCACCACCCAGCTGGGACCTTATGCCTTCAAGATCCCACTGTCCATCCGCCAG 2286

Qy   2541 AAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCC 2600

Db   2287 AAGATATGCAACAGCCTAGATGCCCCCAACTCACGGGGCAATGACTGGCGGATGTTAGCA 2346

Qy   2601 CAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCC 2660

Db   2347 CAGAAGCTCTCTATGGACCGGTACCTGAATTACTTTGCCACCAAAGCGAGCCCCACGGGT 2406

Qy   2661 ATGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCT 2720

Db   2407 GTGATCCTGGACCTCTGGGAAGCTCTGCAGCAGGACGATGGGGACCTCAACAGCCTGGCG 2466

Qy   2721 GCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAG 2780

Db   2467 AGTGCCTTGGAGGAGATGGGCAAGAGTGAGATGCTGGTGGCTGTGGCCACCGACGGGGAC 2526

Qy   2781 TGCTGA 2786

Db   2527 TGCTGA 2532
```

```
RESULT 6
AY406492
LOCUS       AY406492                2802 bp    DNA     linear   GSS 15-DEC-2003
DEFINITION  Pan troglodytes UNC5C gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY406492
VERSION     AY406492.1  GI:39762466
KEYWORDS    GSS.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1  (bases 1 to 2802)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 2802)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1..2802
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
     gene            <1..>2802
                     /gene="UNC5C"
                     /locus_tag="HCM2575"
ORIGIN

  Query Match             26.5%;  Score 763.2;  DB 29;  Length 2802;
```

```
Best Local Similarity   52.9%;  Pred. No. 8e-118;
Matches 1456;  Conservative    0;  Mismatches 1218;  Indels   78;  Gaps    7;

Qy         98 GCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGG 157
Db         66 GCTCGTGCTACCTGCCCTGGCCCTGCTCAGCGCCAGCGGCACTGGCTCCGCCGCCCAAGA 125

Qy        158 TGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCC 217
Db        126 TGATGACTTTTTTCATGAACTCCCAGAAACTTTTCCTTCGGATCCACCTGAGCCTCTGCC 185

Qy        218 CCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGT 277
Db        186 ACATTTCCTTATTGAACCTGAGGAAGCTTATATTGTGAAGAATAAGCCTGTGAACCTGTA 245

Qy        278 GTGCAAGGCC------GTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGT 331
Db        246 CTGTAAAGCAAGCCCTGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 305

Qy        332 GCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCGACCAT 391
Db        306 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 365

Qy        392 GGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATA 451
Db        366 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 425

Qy        452 CTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACAT 511
Db        426 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 485

Qy        512 CCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCT 571
Db        486 NNNNNNNNNNNNNNNNCGGAAGACATTTGAGCAGGAACCCCTAGGAAAGGAAGTGTCTTT 545

Qy        572 GGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGA 631
Db        546 GGAACAGGAAGTCTTACTCCAGTGTCGGCCACCTGAAGGGATCCCAGTGGCTGAGNNNNN 605

Qy        632 GTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCG 691
Db        606 NNNNNNGAAAAATGAAGACATAATTGATCCCGTTGAAGATCGGAATTTTTATATTACTAT 665

Qy        692 GGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGT 751
Db        666 TGATCACAACCTCATCATAAAGCAGGCCCGACTCTCTGATACTGCAAATTATACCTGTGT 725

Qy        752 GGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTACGTGAA 811
Db        726 TGCCAAAAACATTGTTGCCAAGAGGAAAAGTACAACTGCCACTGTCATAGTCTATGTCAA 785

Qy        812 CGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTG 871
Db        786 CGGTGGCTGGTCCACCTGGACGGAGTGGTCTGTGTGTAACAGCCGCTGTGGACGAGGGTA 845

Qy        872 GCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGA 931
Db        846 TCAGAAACGTACAAGGACTTGTACCAACCCGGCACCACTCAATGGGGGTGCCTTCTGTGA 905

Qy        932 GGGGCAGAATGTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGCAGCTG 991
Db        906 AGGGCAGAGTGTGCAGAA---AATAGCCTGTACTACGTTATGCCCAGTGGATGGCAGGTG 962

Qy        992 GAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGA 1051
Db        963 GACGCCATGGAGCAAGTGGTCTACTTGTGGAACTGAGTGCACCCACTGGCGCAGGAGGGA 1022

Qy       1052 GTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACAC 1111
Db       1023 GTGCACGGCGCCAGCCCCCAAGAATGGAGGCAAGGACTGCGATGGCCTCGTCTTGCAATC 1082

Qy       1112 CCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCT 1171
Db       1083 CAAGAACTGCACTGATGGGCTTTGCATGCAGACTGCTCCTGATTCAGATGATGTTGCTCT 1142

Qy       1172 CTATGTGGG---CCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCAT 1228
Db       1143 CTATGTTGGGATTGTGATAGCAGTGATCGTTTGCCTGGCGATCTCTGTAGTTGTGGCCTT 1202

Qy       1229 CCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCT 1288
Db       1203 GTTTGTGTATCGGAAGAATCATCGTGACTTTGAGTCAGATATTATTGACTCTTCGGCACT 1262

Qy       1289 CACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCT 1348
Db       1263 CAATGGGGGCTTTCAGCCTGTGAACATCAAG---------GCAGCAAGACAAGATCTGCT 1313

Qy       1349 CACCATCCAGCCGGACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCA 1408
Db       1314 GGCTGTACCCCCAGACCTCACGTCAGCTGCAGCCATGTACAGAGGACCTGTCTATGCCCT 1373

Qy       1409 GGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGG 1468
Db       1374 GCATGACGTCTCAGACAAAATCCCAATGACCAACTCTCCAATTCTGGATCCACTGCCCAA 1433

Qy       1469 CGGCCGCCACACACTGCACCACAGCT-----------CTCCCACCTCTGAGGCCGAGGA 1516
Db       1434 CCTGAAAATCAAAGTGTACAACACCTCAGGTGCTGTCACCCCCCAAGATGACCTCTCTGA 1493

Qy       1517 GTTCGTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAA 1576
Db       1494 GTTTACGTCCAAGCTGTCCCCTCAGATGACCCAGTCATTGTTGGAGAATGAAGCCCTCAG 1553

Qy       1577 CATGA---------------------------------------CCTATGGGACCTT 1594
Db       1554 CCTGAAGAACCAGAGTCTAGCAAGGCAGACTGATCCATCCTGTACCGCATTTGGCAGCTT 1613

Qy       1595 CAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCC 1654
Db       1614 CNNNTNGCTGGGNNNNCACCTTATTGTTCCCAATTCAGGAGTCAGCTTGCTGATTCCCGC 1673

Qy       1655 AGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGA 1714
Db       1674 TGGGGCCATTCCCCAAGGGAGAGTCTACGAAATGTATGTGACTGTACACAGGAAAGAAAC 1733

Qy       1715 CGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACC 1774
Db       1734 TATGAGGCCACCCATGGATGACTCTCAGACACTTTTGACCCCTGTGGTGAGCTGTGGGCC 1793

Qy       1775 CCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAG 1834
Db       1794 CCCAGGAGCTCTGCTCACCCGCCCCGTCGTCCTTACTATGCATCACTGCGCAGACCCCAA 1853

Qy       1835 CCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGCAGGA 1894
Db       1854 TACCGAGGACTGGAAAATACTGCTCAAGAACCAGGCAGCACAGGGACAGTGGG---AGGA 1910

Qy       1895 TGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAG 1954
Db       1911 TGTGGTGGTGGTCGGGGAGGAAAACTTCACCACCCCCTGCTACATTCAGCTGGATGCAGA 1970

Qy       1955 TGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAG 2014
Db       1971 GGCCTGCCACATCCTCACAGAGAACCTCAGCACCTACGCCCTGGTAGGACATTCCACCAC 2030

Qy       2015 CGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCT 2074
Db       2031 CAAAGCGGCTGCGAAGCGCCTCAAGCTGGCCATCTTTGGGCCCCTGTGCTGCTCCTCGCT 2090

Qy       2075 CGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGT 2134
Db       2091 GGAGTACAGCATCCGAGTCTACTGTCTGGATGACACCCAGGATGCCCTGAAGGAAATTTT 2150

Qy       2135 GCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAA 2194
Db       2151 ACATCTTGAGAGACAGATGGGAGGACAGCTCCTAGAAGAACCTAAGGCTCTTCATTTTAA 2210
```

```
Qy   2195 GGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAG 2254
Db   2211 AGGCAGCACCCACAACCTGCGCCTGTCAATTCACGATATCGCCCATTCCCTCTGGAAGAG 2270

Qy   2255 TAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCG 2314
Db   2271 CAAATTGCTGGCTAAATATCAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2330

Qy   2315 GTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTG 2374
Db   2331 NAACCTGCACTGCACCTTCACTCTGGAAAGATTTAGCCTGAACACAGTGGAGCTGGTTTG 2390

Qy   2375 CAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACAT 2434
Db   2391 CAAACTCTGTGTGCGGCAGGTGGAAGGAGAAGGGCAGATCTTCCAGCTCAACTGCACCGT 2450

Qy   2435 CACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGC 2494
Db   2451 GTCAGAGGAACCTACTGGCATCGATTTGCCGCTGCTGGATCCTGCGAACACCATCACCAC 2510

Qy   2495 CCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAG 2554
Db   2511 GGTCACGGGGCCCAGTGCTTTCAGCATCCCTCTCCCTATCCGGCAGAAGCTCTGTAGCAG 2570

Qy   2555 CCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCT 2614
Db   2571 CCTGGATGCCCCCCAGACGAGAGGCCATGACTGGAGGATGCTGGCCCATAAGCNNNNNNN 2630

Qy   2615 GGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCT 2674
Db   2631 NNNNNNGTACTTGAATTACTTTGCCACCAAATCCAGCCCAACTGGCGTAATCCTGGATCT 2690

Qy   2675 GTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGG 2734
Db   2691 TTGGGAAGCACAGAACTTCCCAGATGGAAACCTGAGCATGCTGGCAGCTGTCTTGGAAGA 2750

Qy   2735 ACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2786
Db   2751 AATGGGAAGACATGAAACGGTGGTGTCCTTAGCAGCAGAAGGGCAGTATTAA 2802
```

```
RESULT 7
BI758231
LOCUS       BI758231                1034 bp    mRNA    linear   EST 25-SEP-2001
DEFINITION  603029876F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5200171 5',
            mRNA sequence.
ACCESSION   BI758231
VERSION     BI758231.1  GI:15749809
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1034)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11501  row: g  column: 20
            High quality sequence stop: 793.
FEATURES             Location/Qualifiers
     source          1..1034
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5200171"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_114"
                     /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: EcoRV (destroyed); RNA source anonymous pool of 6
                     male_brains, age range 23-27 yo. Library is oligo-dT
                     primed and directionally cloned (EcoRV site is  destroyed
                     upon cloning).  Average insert size 1.5 kb, insert size
                     range 1-3 kb.  Library is normalized and enriched for
                     full-length  clones and was constructed by C. Gruber
                     (Invitrogen).  Research  Genetics tracking code 019. Note:
                     this is a NIH_MGC Library."
ORIGIN

  Query Match            26.2%;  Score 756.2;  DB 12;  Length 1034;
  Best Local Similarity  89.7%;  Pred. No. 7.5e-117;
  Matches  905;  Conservative    0;  Mismatches   88;  Indels   16;  Gaps    8;

Qy    191 GCCTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACAT 250
Db      1 GCCTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACAT 60

Qy    251 CGTCAAGAACAAGCCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTT 310
Db     61 CGTCAAGAACAAGCCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTT 120

Qy    311 CAAGTGCAACGGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGG 370
Db    121 CAAGTGCAACGGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGG 180

Qy    371 GAGCAGTGGTGAGCCGACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAA 430
Db    181 GAGCAGTGGGCTGCCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAA 240

Qy    431 GGTGTTCGGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCAC 490
Db    241 GGTGTTCGGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCAC 300

Qy    491 CAAGAGTCAGAAGGCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCC 550
Db    301 CAAGAGTCAGAAGGCCTACATCCGCATAGCCTATTTGCGCAAGAACTTCGAGCAGGAGCC 360

Qy    551 GCTGGCCAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGG 610
Db    361 GCTGGCCAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGG 420

Qy    611 CATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGA 670
Db    421 CATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGA 480

Qy    671 CCCCAATGTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGA 730
Db    481 CCCCAATGTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGA 540

Qy    731 CACGGCCAACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGC 790
Db    541 CACGGACAACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGC 600

Qy    791 TGCTGTCATCGTCTACGTGAACGGTGGGTGGTCGA-CGTGGACCGAGTGGTCCGTCTGCA 849
Db    601 TGCTGTCATCGTCTACGTGAACGGTGGGTGGTCGACCGTGGACCGAGTGGTCCGTCTGC- 659

Qy    850 GCGCCAGCTGTGGGCGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTC 909
Db    660 GCGCCAGCTGTGGGCGCGGCTGGCAGAAACCGGAGCCGGAGCTGCACAACCCGGTGCCTC 719

Qy    910 TCAACGGGGGCGCTTTCTGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCCTCTCTGC 969
Db    720 TCAACGGGGGCGCTTTCTGTGAGGGGGCAGA-----ATGTCCAGAAAGCAGCTGCGCCAC 774

Qy    970 TTGTCTCTGTGGACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGT--GGGCTGGA 1027
```

```
Db        775 CTGTGCCAGTGGACGGCAGCTGTAGCACGTGGAGCCAGTGGTCGGCCTGTTGGGCTTGGC 834

Qy       1028 CTGCACCCACT--GGCGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAG 1085

Db        835 TTGCACCCACTTGGGCGGAGCCGGAGTGCTCTGAACCCAGCACCCCGGCACGGCAGGGGG 894

Qy       1086 GAGTG---CCAGGGCACTGACCTGGACACCCGCAACTGTACCAGTGACCTCTGTGTACAC 1142

Db        895 AGGTGTGCCCAGGTCACTGGACCTGGCACCCGGGA-TGGTCCAGTGAGCTCTGTGT-CCC 952

Qy       1143 AGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGCCTCATCGCCG 1191

Db        953 ACGGGTTCTGGCCCTGAGGACTTGGCCTCCTATGTGGGCCTCATCCCGG 1001


RESULT 8
AY411749
LOCUS       AY411749                2532 bp    DNA     linear   GSS 12-DEC-2003
DEFINITION  Mus musculus HCM4327 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY411749
VERSION     AY411749.1  GI:39767717
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 2532)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 2532)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1..2532
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
     gene            <1..>2532
                     /locus_tag="HCM4327"
ORIGIN

  Query Match             25.2%;  Score 726.4;  DB 29;  Length 2532;
  Best Local Similarity   59.1%;  Pred. No. 1.2e-111;
  Matches 1492;  Conservative    0;  Mismatches  896;  Indels  138;  Gaps    9;

Qy        393 GAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATAC 452

Db         13 GAGGTGCAGATCGAGGTGTCACGGCAGCAAGTGGAGGAACTCTTCGGGCTCGAGGACTAC 72

Qy        453 TGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATC 512

Db         73 TGGTGCCAGTGCGTGGCCTGGAGCTCTTCGGGAACTACCAAGAGTCGCCGAGCCTACATC 132

Qy        513 CGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTG 572

Db        133 CGCATTGCCTACTTGCGCAAGAACTTTGACCAGGAGCCTCTGGCCAAGGAGGTACCCTTG 192

Qy        573 GAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAG 632

Db        193 GATCATGAGGTCCTTCTGCAGTGCCGCCCACCGGAGGGAGTGCCTGTGGCTGAGGTGGAA 252

Qy        633 TGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGG 692

Db        253 TGGCTCAAGAATGAAGATGTCATTGACCCCGCTCAGGACACTAACTTCCTGCTCACCATT 312

Qy        693 GAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTG 752

Db        313 GACCACAACCTCATCATCCGCCAGGCGCGCCTCTCAGACACGGCCAACTACACCTGTGTG 372

Qy        753 GCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTACGTGAAC 812

Db        373 GCCAAGAATATCGTGGCCAAGCGCCGGAGCACCACGGCCACAGTCATCGTCTATGTGAAT 432

Qy        813 GGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGG 872

Db        433 GGAGGCTGGTCCAGCTGGGCAGAGTGGTCACCCTGTTCCAATCGCTGTGGCCGAGGCTGG 492

Qy        873 CAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAG 932

Db        493 CAGAAGCGTACTCGGACCTGCACCAATCCAGCCCCACTCAATGGAGGCGCCTTCTGTGAG 552

Qy        933 GGGCAGAATGTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGCAGCTGG 992

Db        553 GGACAGGCCTTCCAGAA---GACAGCTTGCACCACCGTGTGCCCAGTGGATGGAGCGTGG 609

Qy        993 AGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAG 1052

Db        610 ACCGAGTGGAGCAAGTGGTCTGCCTGCAGCACAGAGTGTGCGCACTGGCGCAGCCGCGAG 669

Qy       1053 TGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACC 1112

Db        670 TGCATGGCACCGCCACCCCAGAACGGAGGCCGTGACTGCAGCGGGACGCTACTTGACTCC 729

Qy       1113 CGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGAC--------- 1163

Db        730 AAGAACTGCACTGATGGGCTGTGCGTGCTGAATCAGAGAACTCTAAACGACCCTAAAAGC 789

Qy       1164 ----------------------GTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTC 1199

Db        790 CACCCCCTGGAGACATCGGGAGATGTGGCACTGTACGCAGGCCTTGTGGTGGCCGTCTTT 849

Qy       1200 TGCCTGGTCCTGCTGCTGCTTGTCCT---CATCCTCGTTTATTGCCGGAAGAAGGAGGGG 1256

Db        850 GTGGTGGTAGCGGTTCTCATGGCCGTGGGAGTGATCGTATACCGGAGAAACTGCCGGGAC 909

Qy       1257 CTGGACTCAGATGTGGCTGACTCGT---CCATTCTCACCTCAGGCTTCCAGCCCGTCAGC 1313

Db        910 TTCGACACGGACATCACCGACTCCTCTGCGGCCCTCACTGGTGGCTTCCACCCTGTCAAC 969

Qy       1314 ATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCT------CACCATCCAGCCGGACCTC 1367

Db        970 TTCAAGACTGCAAGGCCCAACAACCCGCAGCTCCTGCACCCGTCCGCCCCTCCAGACCTA 1029

Qy       1368 AGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCCAGCCCCAAG 1427

Db       1030 ACGGCCAGTGCTGGCATCTACCGCGGGCCTGTGTATGCCCTGCAGGACTCCGCCGACAAG 1089

Qy       1428 TTCCAGCTCACCAATGGGCACCTGCT-------------------------------- 1453

Db       1090 ATCCCCATGACTAATTCGCCCCTGCTGGATCCCCTGCCCAGCCTCAAGATCAAGGTCTAT 1149

Qy       1454 ------------------------------------------CAGCCCCCTGGGTGGC 1469

Db       1150 AACTCCAGCACCATCGGTTCTGGGTCTGGCCTGGCTGATGGAGCCGACCTGCTGGGTGTC 1209

Qy       1470 GGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGC 1529

Db       1210 CTCCCGCCGGGCACGTACCCAGGCGATTTCTCCCGGGACACCCATTTCCTGCACCTGCGC 1269

Qy       1530 CTCTCCACCCAGAACT------ACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC 1583

Db       1270 AGTGCCAGCCTTGGTTCCCAGCACCTCCTGGGCCTACCTCGGGACCCCAGCAGCAGTGTC 1329
```

```
Qy       1584 TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC 1643
Db       1330 AGCGGCACCTTTGGTTGCCTGGGAGGAAGGCTGAGCCTCCCCGGCACAGGGGTCAGCCTG 1389

Qy       1644 CTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC 1703
Db       1390 TTGGTACCAAATGGAGCCATTCCCCAGGGCAAGTTCTATGACCTGTATCTACATATCAAC 1449

Qy       1704 AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT 1763
Db       1450 AAGGCCGAAAGCACCCTCCCACTTTCAGAAGGTTCCCAGACAGTATTGAGCCCCTCGGTG 1509

Qy       1764 AGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT 1823
Db       1510 ACCTGTGGGCCCACAGGCCTACTCCTGTGCCGCCCTGTCGTCCTCACCGTGCCCCACTGT 1569

Qy       1824 GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC 1883
Db       1570 GCTGAAGTCATCGCTGGAGACTGGATCTTTCAGCTCAAGACCCAGGCCCATCAGGGCCAC 1629

Qy       1884 TGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG 1943
Db       1630 TGGG---AGGAGGTGGTGACCTTGGATGAGGAGACCCTCAACACACCCTGCTACTGCCAG 1686

Qy       1944 CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA 2003
Db       1687 CTGGAGGCTAAGTCCTGCCACATCCTGCTGGACCAGCTGGGTACCTACGTATTCATGGGC 1746

Qy       2004 GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC 2063
Db       1747 GAGTCCTACTCTCGCTCTGCAGTCAAGCGGCTCCAGCTGGCCATCTTCGCCCCAGCCCTC 1806

Qy       2064 TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC 2123
Db       1807 TGCACCTCCCTGGAGTATAGCCTCAGGGTCTACTGTCTGGAGGACACACCTGTAGCACTG 1866

Qy       2124 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTC 2183
Db       1867 AAGGAGGTCCTGGAGCTGGAGAGGACTCTGGGTGGCTACTTGGTGGAGGAGCCCAAGCCT 1926

Qy       2184 CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCC 2243
Db       1927 TTGCTCTTTAAGGACAGTTACCACAACCTACGCCTCTCCCTCCATGACATCCCCCATGCC 1986

Qy       2244 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2303
Db       1987 CACTGGAGGAGCAAACTACTGGCCAAGTACCAGGAGATTCCCTTCTACCACGTCTGGAAT 2046

Qy       2304 GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2363
Db       2047 GGCAGCCAGAGAGCCCTGCACTGCACTTTCACCCTGGAGAGGCATAGCCTGGCCTCCACG 2106

Qy       2364 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2423
Db       2107 GAGTTCACCTGTAAGGTCTGCGTGCGGCAGGTCGAAGGGGAAGGCCAGATTTTCCAGCTG 2166

Qy       2424 AACTTCAACATCACC---AAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAA 2480
Db       2167 CACACAACGTTGGCCGAGACGCCTGCTGGCTCCCTGGATGCTCTCTGCTCTGCCCCGGGC 2226

Qy       2481 GCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAG 2540
Db       2227 AATGCCATCACCACCCAGCTGGGACCCTATGCCTTCAAGATACCCCTGTCCATCCGCCAA 2286

Qy       2541 AAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCC 2600
Db       2287 AAGATCTGCAGCAGCCTGGACGCCCCCAACTCCCGGGGCAACGACTGGAGGCTGTTGGCG 2346

Qy       2601 CAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCC 2660
Db       2347 CAGAAGCTGTCCATGGACCGGTACCTAAACTACTTCGCCACCAAAGCTAGTCCCACAGGT 2406

Qy       2661 ATGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCT 2720
Db       2407 GTCATCTTAGACCTCTGGGAAGCTCGGCAACAGGATGACGGGGACCTCAACAGCCTGGCC 2466

Qy       2721 GCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAG 2780
Db       2467 AGTGCCTTGGAGGAGATGGGCAAGAGTGAGATGCTGGTAGCCATGGCCACAGATGGCGAT 2526

Qy       2781 TGCTGA 2786
Db       2527 TGCTGA 2532
```

RESULT 9
BX364574

```
LOCUS       BX364574                 909 bp    mRNA    linear   EST 05-MAY-2003
DEFINITION  BX364574 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
            cDNA clone CS0DC023YB10 5-PRIME, mRNA sequence.
ACCESSION   BX364574
VERSION     BX364574.1  GI:30368812
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 909)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 3529.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS1AC006ZF06QP1&cluster=3529.f. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS1AC006ZF06QP1.
FEATURES             Location/Qualifiers
     source          1..909
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DC023YB10"
                     /tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN

  Query Match             25.2%;  Score 726.2;  DB 13;  Length 909;
  Best Local Similarity   98.4%;  Pred. No. 7.9e-112;
  Matches  739;  Conservative    4;  Mismatches    7;  Indels    1;  Gaps    1;

Qy          1 AGCTGGGGCTCCGGGCTGAGGCGCTAAAGCCGCCCTCCCGCCCGCGGGGCCCCGCGCCCG 60
Db        160 AGCTGGGGCTCCGGGCTGAGGCGCTAAAGCCKCCCTCCCGCCCGCRGGCCCCGCGCCCG 219

Qy         61 GCCCGCCCGCCTGCCCGCCCGCGGCCATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCC 120
Db        220 GCCCGCCCGCCTGCCCGCCCGCGGCCATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCC 279

Qy        121 TGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGG 180
Db        280 TGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGG 339

Qy        181 CCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGG 240
```

```
Db        340 CCAACCCAGTGCCTGGTGCYAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGG 399

Qy        241 ATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        400 ATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGC 459

Qy        301 AGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        460 AGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCA 519

Qy        361 GCACAGACGGGAGCAGTGGTGAGCCGACCATGGAGGTCCGCATTAATGTCTCAAGGCAGC 420
               :||||||||||||||||      |||  ||||||||||||||||||||||||||||||
Db        520 -SACAGACGGGAGCAGTGGGCTGCCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGC 578

Qy        421 AGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 AGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCT 638

Qy        481 CGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCG 540
              ||||||||||||||||||||||||||||||||||||||||    ||||||||||||||||
Db        639 CGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCCTATTTGCGCAAGAACTTCG 698

Qy        541 AGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        699 AGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTC 758

Qy        601 CACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        759 CACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACC 818

Qy        661 CGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        819 CGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCC 878

Qy        721 GCCTTGCTGACACGGCCAACTACACCTGCGT 751
              |||||||||||||||||||||||||||||||
Db        879 GCCTTGCTGACACGGCCAACTACACCTGCGT 909
```

```
RESULT 10
AI951556
LOCUS       AI951556              788 bp    mRNA    linear   EST 09-MAR-2000
DEFINITION  wv36f04.x1 NCI_CGAP_Ov18 Homo sapiens cDNA clone IMAGE:2531647 3'
            similar to TR:O08721 O08721 TRANSMEMBRANE RECEPTOR UNC5H1. ;, mRNA
            sequence.
ACCESSION   AI951556
VERSION     AI951556.1  GI:5743866
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 788)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
            R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: M. Bento
            Soares, Ph.D. cDNA Library Arrayed by: Christa Prange, The
            I.M.A.G.E. Consortium DNA Sequencing by: Washington University
            Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 1125   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 446.
FEATURES             Location/Qualifiers
     source          1..788
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2531647"
                     /tissue_type="fibrotheoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NCI_CGAP_Ov18"
                     /note="Organ: ovary; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGCACATTTTTTTTTTTTTTTTTT 3'];
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia),  digested with Not I and cloned into the Not
                     I and Eco RI sites  of the modified pT7T3 vector. Library
                     went through one round of normalization, and was
                     constructed by Bento Soares and M. Fatima Bonaldo.  "
ORIGIN

  Query Match            25.1%;  Score 722.2;  DB 9;  Length 788;
  Best Local Similarity   96.1%;  Pred. No. 3.5e-111;
  Matches  749;  Conservative    0;  Mismatches   27;  Indels    3;  Gaps    1;

Qy       1510 CCGAGGAGTTCGTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCA 1569
              ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db          1 CCGAGGAGTTCGTCTCCCGCCTCTCCGCCCAGAACTACTTCCGCTCCCTGCCCCGAGGCA 60

Qy       1570 CCAGCAACATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATA 1629
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCAGCAACATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATA 120

Qy       1630 CAGGTATCAGCCTCCTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCT 1689
              |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CAGGAATCAGCCTCCTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCT 180

Qy       1690 ACCTCACGCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGC 1749
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCTCACGCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGC 240

Qy       1750 TGAGTCCCATCGTTAGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGG 1809
              |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db        241 TGAGTCCCATCGTTAGCTGTGGACCCCCTGGCGTTCTGCTCACCCGGCCAGTCATCCTGG 300

Qy       1810 CTATGGACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGT 1869
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTATGGACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGT 360

Qy       1870 CGTGCGAGGGCAGCTGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACC 1929
              ||||||||||||||||   |||||||||||||||||||||||||||||||||||||||||
Db        361 CGTGCGAGGGCAGCTGGG---AGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACC 417

Qy       1930 TCTACTACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCT 1989
              ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        418 TCTACTACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGAGCCGCT 477

Qy       1990 TTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGT 2049
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 ATGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGT 537

Qy       2050 TTGCGCCGGTGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACA 2109
              |||||||||||||||||||||||||||||||||||||| | |||||||||||||||||||
Db        538 TTGCGCCGGTGGCCTGCACCTCCCTCGAGTACAACATACTGGTCTACTGCCTGCATGACA 597

Qy       2110 CCCACGATGCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCC 2169
              | |||||||||||| |  |||||||||||||||||||||||||||  |||||||||||||
Db        598 CTCACGATGCACTCAACGTAGTGGTGCAGCTGGAGAAGCAGCTGCAGGGACAGCTGATCC 657

Qy       2170 AGGAGCCACGGGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACG 2229
              ||||||||  ||| |||||||||||||||||||||||||||||||||||||| |||||||
Db        658 AGGAGCCACTGGTACTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATNCATCCACG 717

Qy       2230 ATGTGCCCAGCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTT 2288
              ||||||||||| ||   |||||||||||| |||||||||||||| |||||| || |
Db        718 ATGTGCCCAGCTNCCNTGTGGAGAGTAAGCTTCTTGTCAGCTACCCAGAGATCCNCTAT 776
```

```
RESULT 11
BC033727
LOCUS       BC033727                1532 bp    mRNA    linear   HTC 19-NOV-2003
DEFINITION  Homo sapiens netrin receptor Unc5h1, mRNA (cDNA clone
            IMAGE:5166762), containing frame-shift errors.
ACCESSION   BC033727
VERSION     BC033727.1  GI:21707230
KEYWORDS    HTC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1532)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  MEDLINE   22388257
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1532)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUL-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC),
            Gaithersburg, Maryland;
            Web site: http://www.nisc.nih.gov/
            Contact: nisc_mgc@nhgri.nih.gov
            Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
            Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
            Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
            Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
            Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D.,McCloskey,J.C.,
            McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
            Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
            Young,A., Zhang,L.-H. and Green,E.D.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 68 Row: i Column: 2
            This clone was selected for full length sequencing because it
            passed the following selection criteria: GenomeScan gene prediction
            This clone has the following problem: frame shifted.
FEATURES             Location/Qualifiers
     source          1..1532
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5166762"
                     /tissue_type="Brain, adult medulla"
                     /clone_lib="NIH_MGC_119"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
ORIGIN

  Query Match             24.7%;  Score 711.2;  DB 11;  Length 1532;
  Best Local Similarity   83.3%;  Pred. No. 3.3e-109;
  Matches  910;  Conservative    0;  Mismatches    8;  Indels  174;  Gaps    2;

Qy        378 GGTGAGCCGACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTC 437
              ||   ||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        612 GGGCTGCCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTC 671

Qy        438 GGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGT 497
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        672 GGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGT 731

Qy        498 CAGAAGGCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCC 557
              ||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
Db        732 CAGAAGGCCTACATCCGCATAGCCTATTTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCC 791

Qy        558 AAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCT 617
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        792 AAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCT 851

Qy        618 CCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAAT 677
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        852 CCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAAT 911

Qy        678 GTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCC 737
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        912 GTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCC 971

Qy        738 AACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTC 797
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        972 AACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTC 1031

Qy        798 ATCGTCTACGTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGC 857
              |||||||||
Db       1032 ATCGTCTAC--------------------------------------------------- 1040

Qy        858 TGTGGGCGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGG 917

Db       1041 ------------------------------------------------------------ 1040

Qy        918 GGCGCTTTCTGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCT 977

Db       1041 ------------------------------------------------------------ 1040

Qy        978 GTGGACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCAC 1037
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1041 GTGGACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCAC 1100

Qy       1038 TGGCGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGC 1097
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1101 TGGCGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGC 1160

Qy       1098 ACTGACCTGGACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCT 1157
              |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db       1161 ACTGACCTGGACACCCGCAACTGTACCAGTGACCTCTGTGTACACACTGCTTCTGGCCCT 1220

Qy       1158 GAGGACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTG 1217
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1221 GAGGACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTG 1280

Qy       1218 CTTGTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGAC 1277
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1281 CTTGTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGAC 1340
```

```
Qy       1278 TCGTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAAC 1337
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1341 TCGTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAAC 1400

Qy       1338 CCCCATCTGCTCACCATCCAGCCGGACCTCAG---CACCACCACCACCTACCAGGGCAGT 1394
              ||||||||||||||||||||||||||||||||   |||||||||||||||||||||||||
Db       1401 CCCCATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGT 1460

Qy       1395 CTCTGTCCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTC 1454
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1461 CTCTGTCCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTC 1520

Qy       1455 AGCCCCCTGGGT 1466
              ||||||||||||
Db       1521 AGCCCCCTGGGT 1532
```

```
RESULT 12
BI818609
LOCUS       BI818609                 818 bp    mRNA    linear   EST 04-OCT-2001
DEFINITION  603033362F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5174559 5',
            mRNA sequence.
ACCESSION   BI818609
VERSION     BI818609.1  GI:15929902
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 818)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11434  row: 1  column: 16
            High quality sequence stop: 744.
FEATURES             Location/Qualifiers
     source          1..818
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5174559"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_115"
                     /note="Organ: pooled brain, lung, testis; Vector:
                     pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
                     source anonymous pool of 6 male brains, age range 23-27; 1
                     male lung, age 27; and 1 male testis, age 69.  Library is
                     oligo-dT primed and directionally cloned (EcoRV site is
                     destroyed upon  cloning).  Average insert size 1.8 kb,
                     insert size range 1-3 kb. Library is normalized and
                     enriched for full-length clones and was  constructed by C.
                     Gruber (Invitrogen).  Research Genetics tracking  code
                     021. Note: this is a NIH_MGC Library."
ORIGIN

  Query Match            24.6%;  Score 710;  DB 12;  Length 818;
  Best Local Similarity   97.1%;  Pred. No. 4e-109;
  Matches  744;  Conservative    0;  Mismatches   20;  Indels    2;  Gaps    2;

Qy         84 GCCATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGG 143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCCATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGG 60

Qy        144 CTCCGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAAC 203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTCCGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAAC 120

Qy        204 CCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAG 263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAG 180

Qy        264 CCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGG 323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGG 240

Qy        324 GAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAG 383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||    | 
Db        241 GAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTG 300

Qy        384 CCGACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTG 443
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTG 360

Qy        444 GAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAG 503
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAG 420

Qy        504 GCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAG 563
              |||||||||||||||||    |||||||||||||||||||||||||||||||||||||||
Db        421 GCCTACATCCGCATAGCCTATTTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAG 480

Qy        564 GTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCC 623
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCC 540

Qy        624 GAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATAC 683
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATAC 600

Qy        684 ATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTAC 743
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTAC 660

Qy        744 ACCTGCGTGGCCAAGAACATCGTGGCACGTC-GCCGCAGCGCCTCCGCTGCTGTCATCGT 802
              ||||||||||||||||||||||||||||||| ||||||| ||||||||||||||||||||
Db        661 ACCTGCGTGGCCAAGAACATCGTGGCACGTCAGCCGCAGGGCCTCCGCTGCTGTCATCGT 720

Qy        803 CTACGTGAACGGTGGGTGGTCGACG-TGGACCGAGTGGTCCGTCTG 847
              |||||| ||||  | |||    || ||||  | ||||||| | |||
Db        721 CTACGTGGACGGCAGCTGGAGCCCGTTGGAGCCAGTGGTCGGGCTG 766
```

```
RESULT 13
BX348193/c
LOCUS       BX348193                 843 bp    mRNA    linear   EST 05-MAY-2003
DEFINITION  BX348193 Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED Homo sapiens
            cDNA clone CS0DB008YE02 5-PRIME, mRNA sequence.
ACCESSION   BX348193
VERSION     BX348193.1  GI:30367258
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 843)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 3239.r For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0BAF004ZD01_AF00293_1&cluster=3239.r.
```

```
                Contact : Feng Liang Email : fliang@lifetech.com URL :
                http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
                Faraday Avenue Genoscope sequence ID : CS0BAF004ZD01_AF00293_1.
FEATURES             Location/Qualifiers
     source          1..843
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DB008YE02"
                     /tissue_type="NEUROBLASTOMA COT 10-NORMALIZED"
                     /clone_lib="Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN

  Query Match           24.6%;  Score 709.8;  DB 13;  Length 843;
  Best Local Similarity  95.0%;  Pred. No. 4.3e-109;
  Matches  776;  Conservative    0;  Mismatches   34;  Indels    7;  Gaps    4;

Qy       1694 CACGCTGCACAAGCCGGAAGACGTGAGG--TTGCCCCTAGCTGGCTGTCAGACCCTGCTG 1751
              ||||| ||  ||  |  |||||||    |||  |  |||  | | ||||||||||||||| ||
Db        824 CACGCGGCCAAAAGCGGAAGAGTGTAGGGTTGCCCCAAACCGGGCTGTCAGACCCTGTTG 765

Qy       1752 AGTCCCATCGTTAGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCT 1811
              |||||||| |||||||||| | |||||||||||||||||||| | ||||||||||||| |
Db        764 AGTCCCATGGTTAGCTGTGAA-CCCCTGGCGTCCTGCTCACCGGCCCAGTCATCCTGGGT 706

Qy       1812 ATGGACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTC-AAAAAGCAGTC 1870
              |||||||||||||||||||||||||||||||||||  ||||||||||| ||||| |||||
Db        705 ATGGACCACTGTGGGGAGCCCAGCCCTGACAGCTGAGGCCTGCGCCTCAAAAAAACAGTC 646

Qy       1871 GTGCGAGGGCAGCTGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCT 1930
              |||||||| ||||||||   || |||| ||||||||||||||| ||||||||||||||||
Db        645 GTGCGAGGACAGCTGGG---AGTATGTCCTGCACCTGGGCGAGNAGGCGCCCTCCCACCT 589

Qy       1931 CTACTACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTT 1990
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        588 CTACTACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTT 529

Qy       1991 TGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTT 2050
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        528 TGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTT 469

Qy       2051 TGCGCCGGTGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACAC 2110
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        468 TGCGCCGGTGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACAC 409

Qy       2111 CCACGATGCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCA 2170
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        408 CCACGATGCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCA 349

Qy       2171 GGAGCCACGGGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGA 2230
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        348 GGAGCCACGGGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGA 289

Qy       2231 TGTGCCCAGCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTA 2290
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        288 TGTGCCCAGCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTA 229

Qy       2291 TCACATCTGGAATGGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAG 2350
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        228 TCACATCTGGAATGGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAG 169

Qy       2351 CCCCAGCACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCA 2410
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        168 CCCCAGCACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCA 109

Qy       2411 GAGCTTCAGCATCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCT 2470
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        108 GAGCTTCAGCATCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCT 49

Qy       2471 GGAGAGTGAAGCGGGGGTCCCAGCCCTGGTGGGCCCC 2507
              ||||||||||||||||||||||| ||  |  ||  |||
Db         48 GGAGAGTGAAGCGGGGGTCCCAGNCCCTGGTGGGCCC 12


RESULT 14
BX345406
LOCUS       BX345406                 939 bp    mRNA    linear   EST 05-MAY-2003
DEFINITION  BX345406 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
            cDNA clone CS0DC023YB10 5-PRIME, mRNA sequence.
ACCESSION   BX345406
VERSION     BX345406.1  GI:30369001
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 939)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 3529.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0AC023DA05QP1&cluster=3529.f. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0AC023DA05QP1.
FEATURES             Location/Qualifiers
     source          1..939
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DC023YB10"
                     /tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN

  Query Match           24.3%;  Score 700.2;  DB 13;  Length 939;
  Best Local Similarity  94.0%;  Pred. No. 1.9e-107;
  Matches  755;  Conservative   22;  Mismatches   20;  Indels    6;  Gaps    5;

Qy          1 AGCTGGGGCTCCGGGCTGAGGCGCTAAAGCCGCCCTCCCGCCCGCGGGGCCCCGCGCCCG 60
              ||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||
Db        142 AGCTGGGGCTCCGGGCTGAGGCGCTAAARCCGCCCTCCCGCCCGCGGGGCCCCGCGCCCG 201

Qy         61 GCCCGCCCGCCTGCCCGCCCGCGGCCATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GCCCGCCCGCCTGCCCGCCCGCGGCCATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCC 261

Qy        121 TGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGG 180
              ||||||||||:|||||||||||||||||||||||||||||||||||||||:|||||||||
Db        262 TGGGCATAGTYCTCGCCGCTTGGCTCCGCGGCTCGGGTGCCCAGCAGAGTKCCACCGTGG 321

Qy        181 CCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGG 240
              ||||||||||:||||||||||||||||||||||||:||||||||||||||||||||||||
Db        322 CCAACCCAGTKCCTGGTGCCAACCCGGACCTGCTTYCCCACTTCCTGGTGGAGCCCGAGG 381

Qy        241 ATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGC 300
              |||||||||| || |||||||||||||| ||| ||:| | |  | | |  |  |||||||||
Db        382 ATGTGTACATTGTTAAGAACAAGCCAGTTCTGTTTKTTTTWAGGTCGTT--CCGCCACGC 439

Qy        301 AGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCA 360
              ||||:|||||:|||||||| | ||||||||||||||||||||||||||||||||||||||
```

```
Db        440 AGATMTTCTTYAAGTGCAA-GTGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCA 498

Qy        361 GCACAGACGGGAGCAGTGGTGAGCCGACCATGGAGGTCCGCATTAATGTCTCAAGGCAGC 420
               ::|:|||||||||:||||   ||| ||||||||||||||||||||||||||||||||||
Db        499 -SMCMGACGGGAGCMGTGGGCTGCCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGC 557

Qy        421 AGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 AGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCT 617

Qy        481 CGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCG 540
              |||||||||||||||||||||||||||||||||||||||||   ||||||:|||:|||||
Db        618 CGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCCTATTTGCGCMAGAMCTTCG 677

Qy        541 AGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        678 AGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTC 737

Qy        601 CACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        738 CACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACC 797

Qy        661 CGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGCCT-GGTGGTGCGACAGGCC 719
              |||||||||||||||||||||||||||||||||::| ||:||| ||||||||||||||||
Db        798 CGTCCCTGGACCCCAATGTATACATCACGCGGGRSCKCASCCTGGGTGGTGCGACAGGCC 857

Qy        720 CGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGC 779
              ||||||||||||||||||:|||||||||||||:|:||||:||||||||||||||||||||
Db        858 CGCCTTGCTGACACGGCCMACTACACCTGCGTKGSCAAGMACATCGTGGCACGTCGCCGC 917

Qy        780 AGCGCCTCCGCTGCTGTCATCGT 802
               |||||||||||||||:|||||
Db        918 -GCGCCTCCGCTGCTGTYATCGT 939
```

RESULT 15
BX345407

```
LOCUS       BX345407                1040 bp    mRNA    linear   EST 05-MAY-2003
DEFINITION  BX345407 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
            cDNA clone CS0DC023YB10 5-PRIME, mRNA sequence.
ACCESSION   BX345407
VERSION     BX345407.1  GI:30369002
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1040)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 3529.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0AC023DA05QP2&cluster=3529.f. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0AC023DA05QP2.
FEATURES             Location/Qualifiers
     source          1..1040
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DC023YB10"
                     /tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN

  Query Match             23.6%;  Score 679.4;  DB 13;  Length 1040;
  Best Local Similarity   94.1%;  Pred. No. 6.2e-104;
  Matches  750;  Conservative    9;  Mismatches   31;  Indels    7;  Gaps    5;

Qy          1 AGCTGGGGCTCCGGGCTGAGGCGCTAAAGCCGCCCTCCCGCCCGCGGGGCCCCGCGCCCG 60
              ||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||
Db        161 AGCTGGGGCTCCGGGCTGAGGCGCTAAARCCGCCCTCCCGCCCGCGGGGCCCCGCGCCCG 220

Qy         61 GCCCGCCCGCCTGCCCGCCCGCGGCCATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 GCCCGCCCGCCTGCCCGCCCGCGGCCATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCC 280

Qy        121 TGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 TGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGG 340

Qy        181 CCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGG 240
              ||||||||||||||||||||||||||||::||||||||||||||||||||||||||||||
Db        341 CCAACCCAGTGCCTGGTGCCAACCCGGAMYTGCTTCCCCACTTCCTGGTGGAGCCCGAGG 400

Qy        241 ATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGC 300
              |||| ||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 ATGTTTACATCKTCAAGAACAAGCCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGC 460

Qy        301 AGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 AGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCA 520

Qy        361 GCACAGACGGGAGCAGTGGTGAGCCGACCATGGAGGTCCGCATTAATGTCTCAAGGCAGC 420
               :|||||||||||||||||   ||| ||||||||||||||||||||||||||||||||||
Db        521 -SACAGACGGGAGCAGTGGGCTGCCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGC 579

Qy        421 AGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        580 AGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCT 639

Qy        481 CGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCG 540
              |||||||||||||||||||||||||||||||||||||||||   ||||||||||||||||
Db        640 CGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCCTATTTGCGCAAGAACTTCG 699

Qy        541 AGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        700 AGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTC 759

Qy        601 CACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        760 CACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACC 819

Qy        661 CGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCC 720
              |||||||||| ||||||||||||||||||||||| :||||||||||||||||||||||||
Db        820 CGTCCCTGGA-CCCAATGTATACATCACGCGGGA-SACAGCCTGGTGGTGCGACAGGCCC 877

Qy        721 GCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCA 780
              | |||||||||||||  |      ||    |||||||||||||||||||||
Db        878 G-CTTGCTGACACGGCAA---CTACACTGCGTGGCAAGAACATCGTGGCACGTCGCGCAS 933

Qy        781 GCGCCTCCGCTGCTGTC 797
              :|  :  | :||   ||
Db        934 SCCTBCGCTSTGTCATC 950
```

Search completed: October  6, 2004, 17:50:24
Job time : 7199 secs